Random Morse Functions

Liviu I. Nicolaescu
University of Notre Dame
Last revised 2025-04-29, 09:35:21.

[1]Started Januray 22, 2024. Completed . Last modified on 2025-04-29, 09:35:21.

Introduction

Morse theory is responsible for many surprising results in geometry and topology. Cleverly chosen Morse functions on concrete manifolds have revealed rich inner structure of those manifolds. Topological applications rely on Morse functions with relatively few critical points. Motivated by questions of the late great V. I. Arnold I became interested in a question that can be loosely formulated as follows: how complex can the Morse functions be?

Notation and conventions

- We set $\mathbb{N} := \mathbb{Z}_{>0}$, $\mathbb{N}_0 := \mathbb{Z}_{\geq 0}$.
- For $n \in \mathbb{N}$ we set $\mathbb{I}_n := \{1, 2, \dots, n\}$.
- For $n \in \mathbb{N}$ we denote by $\mathfrak{S}_n$ the group of permutations of $\mathbb{I}_n$.
- We set $\mathbb{R}_+ := [0, \infty)$.
- For $x \in \mathbb{R}$ we set $\lfloor x \rfloor := \max \mathbb{Z} \cap (-\infty, x]$, $\lceil x \rceil := \min \mathbb{Z} \cap [x, \infty)$.
- $x \wedge y := \min(x, y)$, $x \vee y := \max(x, y)$.
- $\boldsymbol{i} := \sqrt{-1}$
- If $\boldsymbol{X}$ is a finite dimensional Euclidean space, we denote by $\mathbf{Sym}(\boldsymbol{X})$ the space of symmetric operators $\boldsymbol{X} \to \boldsymbol{X}$.
- Given an ambient set Ω and a subset $A \subset \Omega$ we denote by $\boldsymbol{I}_A : \Omega \to \{0, 1\}$ the *indicator function* of A,

$$\boldsymbol{I}_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

- Given a subset A of a set Ω we denote by A^c its complement (in Ω).
- For any set Ω we denote by 2^Ω the collection of all the subsets of Ω.
- For any set Ω we denote by 2_0^Ω the collection of all the *finite* subsets of Ω.
- We will denote by $|S|$ or $\#S$ the cardinality of a set S.
- If T is a topological space, then we denote by $\mathcal{B}_T$ the σ-algebra of Borel subsets of T.
- We denote by $\boldsymbol{\lambda}$ the standard Lebesgue measure on $\mathbb{R}$ and by $\boldsymbol{\lambda}_n$ the standard Lebesgue measure on $\mathbb{R}^n$.

- We denote by $\boldsymbol{\omega}_n$ the volume of the unit ball in $\mathbb{R}^n$ and by $\boldsymbol{\sigma}_{n-1}$ the "area" of the unit ($(n-1)$-dimensional) sphere in $\mathbb{R}^n$.

$$\boldsymbol{\omega}_n = \frac{1}{n}\boldsymbol{\sigma}_{n-1}, \;\; \boldsymbol{\sigma}_{n-1} = \frac{2\Gamma(1/2)^n}{\Gamma(n/2)} = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

The probabilistic notations are those in [**118**].

A *measurable space* is a pair $(\Omega, \mathcal{S})$, where $\mathcal{S}$ a sigma-algebra of subsets of the set Ω. A *measured space* is a triplet $(\Omega, \mathcal{S}, \mu)$, where $(\omega,]\mathcal{S})$ is a measurable space and $\mu : \mathcal{S} \to [0, \infty]$ is a measure.

If $\Phi : (\Omega_0, , \mathcal{S}_0) \to (\Omega_1, \mathcal{S}_1)$ is a measurable map between measurable spaces and μ_0 is a measure on $\mathcal{S}_0$, then the *pushforward* of μ_0 by Φ is the measure $\Phi_\#\mu_0$ on $\mathcal{S}_1$ defined by

$$\Phi_\#\mu_0\big[\, S_1 \,\big] = \mu_0\big[\, \Phi^{-1}(S^1) \,\big], \;\; \forall S_1 \in \mathcal{S}_1.$$

Also, we will often use the notation

$$\big\{\, \Phi \in S_1 \,\big\} := \Phi^{-1}(S_1).$$

The probability spaces are measured spaces $(\Omega, \mathcal{S}, \mathbb{P})$, such that $\mathbb{P}\big[\, \Omega \,\big] = 1$.

Most of the time we will stick to the convention to capitalize the names of random variables. The expectation of a random variable X is denoted by $\mathbb{E}\big[\, X \,\big]$. The conditional expectation of Y given X is denoted by $\mathbb{E}\big[\, Y \,\|\, X \,\big]$. If Y is valued in a finite dimensional vector space $\mathfrak{Y}$ and X in a finite dimensional vector space $\mathfrak{X}$, then there exists a Borel measurable map $F : \mathfrak{X} \to \mathfrak{Y}$ such that $F(X) = \mathbb{E}\big[\, Y \,\|\, X \,\big]$. We set

$$\mathbb{E}\big[\, Y \,\big|\, X = x \,\big] := F(x).$$

The distribution of a random variable $X : (\Omega, \mathcal{S}, \mathbb{P}) \to \mathfrak{X}$, $\mathfrak{X}$ finite dimensional vector space, is denoted by $\mathbb{P}_X$ and it is the pushforward of $\mathbb{P}$ by X, $\mathbb{P}_X = X_\#\mathbb{P}$. More explicitly

$$\mathbb{P}_X\big[\, B \,\big] = \mathbb{P}\big[\, \{X \in B\} \,\big], \;\; \forall B \in \mathcal{B}_{\mathfrak{X}}.$$

Let V be a finite dimensional real vector space. The Euclidean topology on V is the topology defined by a norm on V. Since all the norms on V are equivalent, the Euclidean topology is well defined. Denote by $\mathscr{B}_V$ the sigma-algebra of Borel subsets of V, i.e., the sigma-algebra generated by the subsets open in the Euclidean topology. We denote by $\operatorname{Prob}(V)$ the set of Borel probability measures on V. We let $\langle -, - \rangle$ denote the natural pairing between a vector space and its dual

$$\langle -, - \rangle : V^* \times V \to \mathbb{R}, \;\; \langle \xi, v \rangle = \xi(v).$$

If H is a Hilbert space with inner product $(-, -)_H$ then the *Gramian* matrix determined by the vectors $x_1, \dots, x_N \in H$ is the $N \times N$ matrix

$$G(x_1, \dots, x_N) = \big(\, (x_i, x_j)_H \,\big)_{1 \le i,j \le N}.$$

Given a topological space X and a vector space V we denote by $\underline{V}_X$ the product bundle over X with fiber V.

$$\underline{V}_X = \big(\, V \times X \to X \,\big).$$

Contents

Chapter 1

Gaussian measures and Gaussian fields

The random functions we consider in this book are Gaussian. Most of the time they are described by Fourier/eigenfunction series with coefficients independent normal variables. Since we are interested mostly in geometric questions it is important to have coordinate free description of basic facts of Gaussian analysis.

1.1. Gaussian measures

1.1.1. Finite dimensional Gaussian measures and vectors.

Definition 1.1.1. Let $m \in \mathbb{R}$ and any $v > 0$. The *Gaussian measure* on $\mathbb{R}$ with mean m and variance v is the measure $\mathbf{\Gamma} = \mathbf{\Gamma}_{m,v} \in \operatorname{Prob}(\mathbb{R})$ given by

$$\mathbf{\Gamma}_{m,v}\big[\,dx\,\big] = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2v}} \boldsymbol{\gamma}_{m,v}(x) \boldsymbol{\lambda}\big[\,dx\,\big], \;\; \mathbf{\Gamma}_{m,v}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2v}},$$

where $\boldsymbol{\lambda}$ denotes the Lebesgue measure on $\mathbb{R}$. □

It is not hard to observe that, as $v \searrow 0$, the measure $\mathbf{\Gamma}_{m,v}$ converges weakly to δ_m, the Dirac measure concentrated at m. For this reason we will refer to δ_m as the Gaussian measure of mean m and variance 0, and we set $\mathbf{\Gamma}_{m,0} := \delta_m$. Observe that for any $m \in \mathbb{R}$ and any $v \geq 0$ we have

$$\int_{\mathbb{R}} x \mathbf{\Gamma}_{m,v}\big[\,dx\,\big] = m, \;\; \int_{\mathbb{R}} (x-m)^2 \mathbf{\Gamma}_{m,v}\big[\,dx\,\big] = v.$$

A Gaussian measure on $\mathbb{R}$ is called *nondegenerate* if it is absolutely continuous with respect to the Lebesgue measure $\boldsymbol{\lambda}$. Equivalently, the Gaussian measure is nondegenerate iff its variance is nonzero. For $v \geq 0$ we set

$$\mathbf{\Gamma}_v := \mathbf{\Gamma}_{m=0,v}.$$

For any $c \in \mathbb{R}$ we denote by $\mathcal{R}_c$ the rescaling map $\mathcal{R}_c : \mathbb{R} \to \mathbb{R}$, $\mathcal{R}_c(x) = cx$. A simple computation shows that

$$(\mathcal{R}_c)_\# \mathbf{\Gamma}_v = \mathbf{\Gamma}_{c^2 v}, \;\; \forall c \in \mathbb{R}, \;, v \geq 0. \tag{1.1.1}$$

The measure $\boldsymbol{\Gamma}_1$ is called the *canonical Gaussian measure* on $\mathbb{R}$. The ratio

$$\frac{\boldsymbol{\Gamma}_1\big[\,(x,\infty)\,\big]}{\boldsymbol{\gamma}_1(x)},\;\; x>0$$

is called the Mills ratio and it satisfies the *Mills ratio* inequalities [19]

$$\frac{x}{x^2+1}\boldsymbol{\gamma}_1(x)\leq \boldsymbol{\Gamma}_1\big[\,(x,\infty)\,\big]\leq \frac{1}{x}\boldsymbol{\gamma}_1(x),\;\; \forall x>0. \tag{1.1.2}$$

The Fourier transform of a Borel probability measure μ on $\mathbb{R}$ is the function

$$\widehat{\mu}:\mathbb{R}\to\mathbb{C},\;\; \widehat{\mu}(t)=\int_{\mathbb{R}} e^{\boldsymbol{i}tx}\mu\big[\,dx\,\big].$$

Lévy's theorem shows that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of probability measures on $\mathbb{R}$ converges weakly to a probability measure μ if and only if

$$\lim_{n\to\infty}\widehat{\mu}_n(t)=\widehat{\mu}(t),\;\; \forall t\in\mathbb{R}.$$

We have

$$\widehat{\boldsymbol{\Gamma}}_{m,v}(t)=e^{\boldsymbol{i}tm-vt^2/2},\;\; \forall t\in\mathbb{R}. \tag{1.1.3}$$

Proposition 1.1.2. *Suppose that the sequence $(\mu_n)_{n\in\mathbb{N}}$ of Gaussian measures on $\mathbb{R}$ converges weakly to a probability measure μ. Then μ is Gaussian and*

$$\lim_{n\to\infty} m[\mu_n]=m[\mu],\;\; \lim_{n\to\infty} v[\mu_n]=v[\mu]$$

where $m[-]$ and $v[-]$ denote the mean and respectively the variance of a Gaussian measure.

Proof. Set $m_n=m[\mu_n]$, $v_n=v[\mu_n]$. Let us first prove that the sequence $(m_n)_{n\in\mathbb{N}}$ is bounded.

We argue by contradiction. Then a subsequence converges to $\pm\infty$. For simplicity we assume $m_n\to\infty$. (The case $m_n\to-\infty$ is dealt with in a similar fashion.). Then

$$\mu_n\big[\,(-\infty,m_n]\,\big]=\frac{1}{2},\;\; \forall n.$$

Fix $c\in\mathbb{R}$ such that

$$\mu\big[\,(-\infty,c]\,\big]>\frac{1}{2}\;\text{ and }\;\mu_n\big[\,(-\infty,c]\,\big]\to\mu\big[\,(-\infty,c]\,\big]. \tag{1.1.4}$$

There exists $N=N(c)$ such that, $\forall n\geq N$ $m_n>c$. We deduce that for any $n>N(c)$.

$$\frac{1}{2}=\mu_n\big[\,(-\infty,m_n]\,\big]\geq\mu_n\big[\,(-\infty,c]\,\big].$$

Letting $n\to\infty$ we deduce

$$\mu\big[\,(-\infty,c]\,\big]=\lim_{n\to\infty}\mu_n\big[\,(-\infty,c]\,\big]\leq\frac{1}{2}.$$

This contradicts the choice (1.1.4). Hence the sequence (m_n) is bounded.

Next, we prove that the sequence v_n is also bounded. Indeed, if it were not bounded, then $\limsup v_n=\infty$. Observe that for any $a<b$ we have

$$\mu\big[\,(a,b)\,\big]=\frac{1}{\sqrt{2\pi v_n}}\int_a^b e^{-\frac{(x-m_n)^2}{2v_n}}dx\leq\frac{1}{\sqrt{2\pi v_n}}\int_a^b dx\leq\frac{(b-a)}{\sqrt{2\pi v_n}}.$$

The Portmanteau Theorem implies that $\forall a < b$

$$0 \leq \mu\big[\,(a,b)\,\big] \leq \liminf_{n\to\infty} \mu_n\big[\,(a,b)\,\big] \leq \liminf_{n\to\infty} \frac{(b-a)}{\sqrt{2\pi v_n}} = 0.$$

This is impossible.

Hence on a subsequence n_k we have $m_{n_k} \to m \in \mathbb{R}$, $v_{n_k} \to v \in [0,\infty)$. Hence, $\forall t \in \mathbb{R}$,

$$\widehat{\mu}(t)(\xi) = \lim_{k\to\infty} \widehat{\mu}_{n_k}(t) \overset{(1.1.3)}{=} \lim_{k\to\infty} e^{\boldsymbol{i} m_{n_k} t - \frac{v_{n_k} t^2}{2}} = e^{\boldsymbol{i} m t - \frac{v t^2}{2}}.$$

Hence $\widehat{\mu} = \widehat{\boldsymbol{\Gamma}}_{m,v}$. This proves that μ is also Gaussian. Moreover, we proved that any subsequence of (μ_n) contains a sub-subsequence (μ_{n_j}) such that $m[\mu_{n_j}]$ and $v[\mu_{n_j}]$ converge to m_∞ and respectively v_∞. Since μ_{n_j} converges weakly to $mu = \boldsymbol{\Gamma}_{m,v}$ we deduce $\boldsymbol{\Gamma}_{m,b} = \boldsymbol{\Gamma}_{m_\infty, v_\infty}$. This proves

$$m[\mu_n] \to m[\mu], \;\; v[\mu_n] \to v[\mu].$$

□

Proposition 1.1.3. *The space of polynomials in one variable with real coefficients is dense in* $L^2(\mathbb{R}, \boldsymbol{\Gamma})$

Proof. We follow the elegant argument in [**94**, Sec.V.1.3]. It suffices to show that if $f \in L^2(\mathbb{R}, \boldsymbol{\Gamma})$ and

$$\int_{\mathbb{R}} f(x) x^n \boldsymbol{\Gamma}\big[\,dx\,\big] = 0, \;\; \forall n = 0,1,2,\dots$$

then $f = 0$ $\boldsymbol{\Gamma}$-a.s.. Observe that

$$\int_{\mathbb{R}} |x|^\alpha e^{tx} \boldsymbol{\Gamma}\big[\,dx\,\big] < \infty, \;\; \forall t \in \mathbb{R}, \;\; \forall n \geq 0.$$

Since $\big(\,|x|^\alpha e^{tx}\,\big)^2 = |x|^{2\alpha} e^{2tx}$, we deduce that for any $t \in \mathbb{R}$ any $\alpha \geq 0$ the function $x \mapsto |x|^\alpha e^{tx}$ is in $L^2(\mathbb{R}, \boldsymbol{\Gamma})$.

For $z = t + \boldsymbol{i}s \in \mathbb{C}$ we set

$$F(z) := \int_{\mathbb{R}} e^{\boldsymbol{i}zx} f(x) \boldsymbol{\Gamma}\big[\,dx\,\big].$$

The above discussion shows that $F(z)$ is well defined and $z \mapsto F(z)$ is an entire function. Moreover

$$F^{(n)}(0) = \boldsymbol{i}^n \int_{\mathbb{R}} f(x) x^n \boldsymbol{\Gamma}\big[\,dx\,\big] = 0, \;\; \forall n = 0,1,\dots\,.$$

We deduce from unique continuation that F is identically zero. In turn, this implies that f is a.s. 0 since an L^2 function is uniquely determined by. □

Proposition 1.1.4 (Gaussian integration by parts)**.** *Suppose that* $f, g \in C^1(\mathbb{R})$ *and there exists* $p > 0$ *such that*

$$\sup_{x\in\mathbb{R}} \big(\,|f'(x)| + |g'(x)|\,\big)\big(\,1 + |x|\,\big)^{-p} < \infty.$$

Then

$$\int_{\mathbb{R}} f'(x) g(x) d\boldsymbol{\Gamma}\big[\,dx\,\big] = \int_{\mathbb{R}} f(x)\big(\,-g'(x) + x g(x)\,\big) \boldsymbol{\Gamma}_1\big[\,dx\,\big]. \tag{1.1.5}$$

Proof. For any $L>0$ we have

$$\int_{-L}^{L} f'(x)g(x)\boldsymbol{\Gamma}\big[\,dx\,\big] = \frac{1}{\sqrt{2\pi}}\int_{-L}^{L} f'(x)g(x)e^{-x^2/2}dx$$

(integrate by parts)

$$= \frac{1}{\sqrt{2\pi}} f'(x)g(x)e^{-x^2/2}\Big|_{x=-L}^{x=L} - \frac{1}{\sqrt{2\pi}}\int_{-L}^{L} f(x)\frac{d}{dx}\big(\,g(x)e^{-x^2/2}\,\big)dx$$

$$= \frac{1}{\sqrt{2\pi}} f'(x)g(x)e^{-x^2/2}\Big|_{x=-L}^{x=L} - \frac{1}{\sqrt{2\pi}}\int_{-L}^{L} f(x)\big(\,g'(x) - xg(x)\,\big)e^{-x^2/2}$$

$$= \frac{1}{\sqrt{2\pi}} f'(x)g(x)e^{-x^2/2}\Big|_{x=-L}^{L} - \int_{-L}^{x=L} f(x)\big(\,g'(x) - xg(x)\,\big)\boldsymbol{\Gamma}_1\big[\,dx\,\big].$$

The equality (1.1.5) follows by letting $L\to\infty$. □

Definition 1.1.5. Suppose that $(\Omega,\mathcal{S},\mathbb{P})$ is a probability space. A random variable

$$X:(\Omega,\mathcal{S},\mathbb{P})\to\mathbb{R}$$

is called *Gaussian* if its distribution $\mathbb{P}_X\in\operatorname{Prob}(\mathbb{R})$ is a Gaussian measure. Note that in this case the mean and the variance of X coincide with the mean and variance of $\mathbb{P}_X$. The random variable is called *centered* iff it has mean 0. It is called *nondegenerate* if its variance is nonzero. □

If X is a Gaussian random variable with mean m and variance v, then its *characteristic function* has the description.

$$\Phi_X(t) := \mathbb{E}\big[\,e^{\boldsymbol{i}tX}\,\big] = \widehat{\mathbb{P}}_X(t) = e^{\boldsymbol{i}mt-\frac{vt^2}{2}}.$$

In particular, this shows that the distribution of a Gaussian random variable is uniquely determined by its mean and variance.

Remark 1.1.6. Note that if X,Y are independent Gaussian variables and $a,b\in\mathbb{R}$, then $aX+bY$ is also Gaussian since

$$\Phi_{AX+bY}(t) = \Phi_{aX}(t)\Phi_{bY}(t) = e^{\boldsymbol{i}am_X t-\frac{a^2v_X t^2}{2}}e^{\boldsymbol{i}m_Y t-\frac{b^2v_Y t^2}{2}}$$

$$= \exp\left(\boldsymbol{i}(am_X+bm_y)t - \frac{(a^2v_X^2+b^2v_Y)t^2}{2}\right).$$

□

The following result is a direct consequence of Proposition 1.1.2.

Proposition 1.1.7. *Suppose $(X_n)_{n\in\mathbb{N}}$ is a sequence of Gaussian random variables defined on the same probability space, and X_n converges in distribution to a random variable X. Then X is also a Gaussian random variable and*

$$\mathbb{E}\big[\,X_n\,\big]\to\mathbb{E}\big[\,X\,\big],\;\;\operatorname{Var}\big[\,X_n\,\big]\to\operatorname{Var}\big[\,X\,\big].$$

□

Definition 1.1.5 has one æsthetic deficiency: it is not quite "coordinate free". The next results addresses this issue.

Theorem 1.1.8 (G. Polya)**.** *Suppose that X is a random variable. The following are equivalent.*

(i) *The random variable X is centered Gaussian.*

(ii) *If X, Y are* i.i.d.*, then X and $\frac{1}{\sqrt{2}}\big(X+Y\big)$ have the same distribution.*

(iii) *If X, Y are* i.i.d.*, then for any $\theta \in [0, 2\pi]$, the random variables $(\cos\theta)X + (\sin\theta)Y$ have a distribution independent of θ.*

□

The implications (i) $\implies$ (iii) $\implies$ (ii) immediate. The tricky implication is (ii) $\implies$ (i). For a proof we refer to [**30**, Thm. 3.1] or [**148**, Thm. 2.2.3].

The next characterization highlights the close connection between the concepts of Gaussian random variables and the concept of independence. For a proof we refer to [**59**, Sec.XV.8].

Theorem 1.1.9 (Bernstein)**.** *Suppose that X, Y are* independent *random variables. The following are equivalent.*

(i) *The variables X, Y are Gaussian.*

(ii) *The variables $X+Y$ and $X-Y$ are independent.*

□

Corollary 1.1.10. *Let X be a a random variable and Y an independent copy of it. Then the following are equivalent.*

(i) *The random variable X is centered Gaussian.*

(ii) *The random vectors (X, Y) and $\big(\frac{1}{\sqrt{2}}(X+Y), \frac{1}{\sqrt{2}}(X-Y)\big)$ have identical distributions.*

□

Proposition 1.1.11. *Suppose that X is a centered Gaussian random variable with variance $v = \mathbb{E}\big[X^2\big]$. Then the following hold.*

$$\forall p \in [1,\infty)\ \ \mathbb{E}\big[\,|X|^p\,\big] = \frac{(2v)^{p/2}}{\sqrt{\pi}}\Gamma\Big(\frac{p+1}{2}\Big), \tag{1.1.6}$$

$$\mathbb{E}\big[\,e^{tX}\,\big] = e^{vt^2/2}. \tag{1.1.7}$$

In particular,

$$\mathbb{E}\big[\,\big|\,X\,\big|\,\big] = (2v/\pi)^{1/2}, \tag{1.1.8}$$

$$\forall k \in \mathbb{N},\ \ \mathbb{E}\big[\,X^{2k-1}\,\big] = 0,\ \ \mathbb{E}\big[\,X^{2k}\,\big] = v^k(2k-1)!!, \tag{1.1.9}$$

where $(2k-1)!! := 1\cdot 3\cdots(2k-1)$.

Proof. Set $Y := v^{-1/2}X$. Then $\mathrm{Var}\big[\,Y\,\big] = 1$ and

$$\mathbb{E}\big[\,|X|^p\,\big] = v^{p/2}\mathbb{E}\big[\,|Y|^p\,\big],\ \ \mathbb{E}\big[\,e^{tX}\,\big] = \mathbb{E}\big[\,e^{\sqrt{v}tY}\,\big].$$

We have

$$\mathbb{E}\big[\,|Y|^p\,\big] = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} |y|^p e^{-y^2/2} dy = \frac{2}{\sqrt{2\pi}}\int_0^\infty |y|^p e^{-y^2/2} dy$$

$(r = y^2/2,\ y = \sqrt{2r})$

$$= \frac{2}{\sqrt{2\pi}}\int_0^\infty (2r)^{p/2} e^{-r}\frac{dr}{(2r)^{1/2}} = \frac{2}{\sqrt{2\pi}}\int_0^\infty (2r)^{\frac{p-1}{2}} e^{-r} dr$$

$$= \frac{2^{(p+1)/2}}{\sqrt{2\pi}}\int_0^\infty r^{\frac{p+1}{2}-1} e^{-r} dr = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\Big(\frac{p+1}{2}\Big).$$

For $p = 1$ we have

$$\mathbb{E}\big[\,|\,Y\,|\,\big] = \sqrt{\frac{2}{\pi}}\Gamma(1) = (2/\pi)^{1/2}.$$

Since the distribution of X is symmetric we deduce

$$\mathbb{E}\big[\,X^{2k-1}\,\big] = 0,\ \ \forall k \in \mathbb{N}.$$

On the other hand,

$$\sum_{n\geq 0}\frac{t^n}{n!}\mathbb{E}\big[\,Y^n\,\big] = \mathbb{E}\big[\,e^{tY}\,\big] = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{ty-y^2/2} dy = \frac{e^{t^2/2}}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{(y-t)^2}{2}} dy = e^{t^2/2}.$$

Hence

$$\mathbb{E}\big[\,Y^{2k}\,\big] = \frac{(2k)!}{2^k k!} = (2k-1)!!.$$

□

Definition 1.1.12. Let V be a finite dimensional real vector space space. A Borel probability measure $\mu \in \mathrm{Prob}(V)$ is called *Gaussian* if for every linear functional $\xi \in V^*$, the induced random variable $\xi : (V, \mathcal{B}_V, \mu) \to \mathbb{R}$ is Gaussian with mean $m_\mu[\xi]$ and variance $v_\mu[\xi]$, i.e.,

$$\mathbb{P}_\xi\big[\,dx\,\big] = \boldsymbol{\Gamma}_{m_\mu[\xi],v_\mu[\xi]}\big[\,dx\,\big].$$

The Gaussian measure μ is called *centered* if $m_\mu[\xi] = 0$, $\forall \xi \in V^*$. We denote by $\mathrm{Gauss}(V)$ the set of Gaussian measures on V and by $\mathrm{Gauss}_0(V)$ the subset of centered ones. □

Example 1.1.13. The Borel measure on $\mathbb{R}^n$ given measure

$$\Gamma_{\mathbb{1}_n} := \boldsymbol{\Gamma}_{0,1}\otimes \dots \otimes \boldsymbol{\Gamma}_{0,1} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\|x\|^2}{2}}\boldsymbol{\lambda}\big[\,dx\,\big],\ \ \|x\|^2 = \sum_{k=1}^n x_k^2,$$

is a Gaussian measure called the *canonical Gaussian measure*. Above $\boldsymbol{\lambda}\big[\,-\,\big]$ denotes the canonical Lebesgue measure on $\mathbb{R}^n$.

To see this consider the coordinate maps

$$X_1, \dots, X_n : \mathbb{R}^n \to \mathbb{R},\ \ X_i(x_1, \dots x_i) = x_i.$$

We view the maps X_i as a random variable defined on the probability space $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \Gamma_{\mathbb{1}_n})$. The measure $\Gamma_{\mathbb{1}_n}$ is the joint distribution of $X_1, \dots, X_n$. In particular, these random variables are independent standard Gaussian with mean 0 and variance one. Using Remark 1.1.6 we deduce that for any $\xi \in \mathbb{R}^n$ the random variable $\xi_1 X_1 + \cdots + \xi_n X_n$ is also Gaussian.

More generally, if $\boldsymbol{U}$ is finite dimensional real Euclidean space with inner product $(-,-)$ and associated norm $\|-\|$, then

$$\Gamma_{\boldsymbol{U}}\big[\,d\boldsymbol{u}\,\big] = (2\pi)^{-\frac{\dim \boldsymbol{U}}{2}} e^{-\frac{1}{2}\|\boldsymbol{u}\|^2} \boldsymbol{\lambda}_{\boldsymbol{U}}\big[\,d\boldsymbol{u}\,\big]$$

is a Gaussian measure called the *canonical Gaussian measure* associated to the metric. Above $\boldsymbol{\lambda}_{\boldsymbol{U}}$ denotes the natural Lebesgue measure on $\boldsymbol{U}$. More precisely,

$$\boldsymbol{\lambda}_{\boldsymbol{U}} = T_{\#}\boldsymbol{\lambda}$$

where $T : \mathbb{R}^{\dim \boldsymbol{U}} \to \boldsymbol{U}$ is any isometry.

To see this fix orthonormal coordinates $u_1, \dots, u_n$ on $\boldsymbol{U}$, $n = \dim \boldsymbol{U}$. In these coordinates

$$\Gamma\big[\,d\boldsymbol{u}\,\big] = (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}(u_1^2+\cdots+u_n^2)} \boldsymbol{\lambda}\big[\,du_1\cdots du_n\,\big] = \Gamma_{\mathbb{1}_n}\big[\,du_1\cdots du_n\,\big].$$

□

Let μ be a Gaussian measure on V. Note that the map $V^* \ni \xi \to m_\mu[\xi] \in \mathbb{R}$ is linear and hence it defines an element in the bidual

$$m_\mu \in V^{**} := \mathrm{Hom}(V^*, \mathbb{R}) = \mathrm{Hom}\big(\,V^*, \mathbb{R}\,\big).$$

Since V is finite dimensional, the natural map $J : V \to V^{**}$ is an isomorphism we can identify $m_\mu \in V^{**}$ with an element of V determined by by the equalities

$$\langle \xi, m_\mu \rangle := m_\mu[\xi], \quad \forall, \xi \in V^*.$$

Equivalently,

$$m_\mu = \int_V v\mu\big[\,dv\,\big].$$

Define the *covariance form* C_μ of the Gaussian measure μ on V to be the map

$$C_\mu : V^* \times V^* \to \mathbb{R}, \quad (\xi, \eta) \mapsto \mathrm{Cov}\big[\,\xi, \eta\,\big] := \mathbb{E}_\mu\big[\,(\xi - m[\xi])(\eta - m[\eta])\,\big].$$

It is not hard to see C_μ is a nonnegative definite symmetric bilinear operator on V^*.

Definition 1.1.14. The *Fourier transform* of a measure $\mu \in \mathrm{Prob}(V)$ is the function

$$\widehat{\mu} : V^* \to \mathbb{C}, \;\; \widehat{\mu}(\xi) = \mathbb{E}_\mu\big[\,e^{\boldsymbol{i}\xi}\,\big] = \int_V e^{\boldsymbol{i}<\xi,x>}\mu[dx]$$

□

Proposition 1.1.15. *A Borel measure μ on V is a Gaussian measure on V if and only if there exists a vector $m \in V$ and a symmetric nonnegative definite bilinear function $C : V^* \times V^* \to \mathbb{R}$ such that*

$$\widehat{\mu}(\xi) = e^{\boldsymbol{i}\langle \xi, m\rangle - \frac{1}{2}C(\xi,\xi)}, \;\; \forall \xi \in V^*.$$

Proof. Indeed, suppose that μ is Gaussian. Then, for any $t \in \mathbb{R}$, $\xi \in V^*$,

$$\widehat{\mu}(t\xi) = \mathbb{E}\big[\,e^{\boldsymbol{i}t\xi}\,\big] = \Phi_\xi(t) = e^{\boldsymbol{i}tm_\mu[\xi] - \frac{t^2 v_\mu[\xi]}{2}},$$

Letting $t = 1$, we deduce

$$\widehat{\mu}(\xi) = e^{\boldsymbol{i}m_\mu[\xi] - \frac{1}{2}C_\mu(\xi,\xi)}.$$

Conversely, if

$$\widehat{\mu}(\xi) = e^{\boldsymbol{i}\langle \xi, m\rangle - \frac{1}{2}C(\xi,\xi)},$$

then for any ξ in V^* and $t \in \mathbb{R}$ we have

$$\Phi_\xi(t) = \hat{\mu}(t\xi) = e^{\boldsymbol{i}t\langle \xi, m\rangle - \frac{t^2 C(\xi,\xi)}{2}},$$

proving that ξ is Gaussian with mean $\langle \xi, m\rangle$ and variance $C(\xi,\xi)$. □

Corollary 1.1.16. *A Gaussian measure on V is uniquely determined by its mean and covariance. Hence, we denote by $\boldsymbol{\Gamma}_{m,C}$ the Gaussian measure with mean m and covariance C.*

Proof. Proposition 1.1.15 shows that the Fourier transform of a Gaussian measure is uniquely determined by the mean and covariance, while the measure is uniquely determined by its Fourier transform. □

Suppose that the vector space V is equipped with an inner product $(-,-)$. The inner product induces an isomorphism

$$^{\downarrow}: V \to V^*,\ \ v \mapsto v^{\downarrow},\ \ \langle v^{\downarrow}, u\rangle = (v,u),\ \ \forall u, v \in V.$$

Classically, this isomorphism is referred to as *lowering the indices*. Its inverse is given by

$$_{\uparrow}: V^* \to V,\ \ (\xi_{\uparrow}, u) = \langle \xi, v\rangle,\ \ \forall \xi \in V^*,\ \ \in V.$$

and it is classically referred to as *raising the indices*.

If μ is a Gaussian measure on V, then its covariance form

$$C_\mu : V^* \times V^* \to \mathbb{R}$$

can be identified with a selfadjoint operator $\operatorname{Var}_\mu : V \to V$ uniquely determined by the equality

$$(u, \operatorname{Var}_\mu v) = C_\mu(u^{\downarrow}, v^{\downarrow}).$$

We will refer to Var_μ as the *variance* (operator) of the measure μ.

Concretely, if (e_i) is a basis of V *orthonormal* with respect to the inner product $(-,-)$, then $X = \sum_i X_i e_i$ an Var_μ is described in this basis by the symmetric matrix (v_{ij}) where

$$v_{ij} = C_\mu(X_i, X_j) = \operatorname{Cov}\big[\, X_i, X_j \,\big].$$

Note that the variance operator of the canonical Gaussian measure on V is $\mathbb{1}_V$.

Remark 1.1.17. The variance operator defined above *depends* on the choice of inner product whereas the covariance form does not. This aspect is important in geometric applications and we want to discuss it in some details.

Let μ be a centered Gaussian measure on the real vector space $\boldsymbol{U}$ of dimension N. Fix two inner products on $\boldsymbol{U}$,

$$(-,-)_i : \boldsymbol{U} \times \boldsymbol{U} \to \mathbb{R},\ \ i = 0, 1.$$

We denote by Var^i the variance operator of μ defined in terms of the inner product $(-,-)_i$. We want to relate Var^0 and Var^1.

Fix an isometry of Euclidean spaces $T : \big(\boldsymbol{U}, (-,-)_1\big) \to \big(\boldsymbol{U}, (-,-)_0\big)$ and set $G = T^*T$. Then $\big(T\boldsymbol{u}, T\boldsymbol{v}\big)_0 = \big(\boldsymbol{u}, \boldsymbol{v}\big)_1$ and $G : \boldsymbol{U} \to \boldsymbol{U}$ is the unique operator that is symmetric and positive definite with respect to inner product $(-,-)_0$ and satisfies

$$(\boldsymbol{u}, \boldsymbol{v})_1 = (G\boldsymbol{u}, \boldsymbol{v})_0,\ \ (\boldsymbol{u}, \boldsymbol{v})_0 = \big(G^{-1}\boldsymbol{u}, \boldsymbol{v}\big)_1,\ \ \forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}.$$

Then for any $\boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{U}$ we have

$$\big(\boldsymbol{v}, \operatorname{Var}^1 \boldsymbol{w}\big)_1 = \int_{\boldsymbol{U}} \big(\boldsymbol{v}, \boldsymbol{u}\big)_1 \big(\boldsymbol{w}, \boldsymbol{u}\big)_1 \mu\big[\boldsymbol{u}\big]$$

$$= \int_{\boldsymbol{U}} \big(G\boldsymbol{v}, \boldsymbol{u}\big)_0 \big(G\boldsymbol{w}, \boldsymbol{u}\big)_0 \mu\big[\boldsymbol{u}\big] = \big(G\boldsymbol{v}, \operatorname{Var}^0 G\boldsymbol{w}\big)_0 = \big(\boldsymbol{v}, \operatorname{Var}^0 G\boldsymbol{w}\big)_1.$$

We deduce that $\operatorname{Var}^1 = \operatorname{Var}^0 G$. □

Let V be a finite dimensional vector space. For $v_0 \in V$ we denote $\mathcal{T}_{v_0}$ the translation operator

$$\mathcal{T}_{v_0} : V \to V, \;\; v \mapsto v + v_0.$$

For any Gaussian measure $\mu \in \mathscr{G}(V)$ the pushforward $(\mathcal{T}_{v_0})_{\#}\mu$ is a Gaussian measure with mean $\mathcal{T}_{v_0} m_\mu = m_\mu + v_0$.

Suppose that V_0 and V_1 are two finite dimensional vector spaces and μ_0 is a Gaussian probability measure on V_0. If $A : V_0 \to V_1$ is a linear map, then we see that the pushforward measure $A_{\#}\mu_0 =: \mu_A$ is Gaussian on V_1 with mean $m_{\mu_A} = Am_\mu$ and covariance form

$$C_{\mu_A} : V^* \times V^* \to \mathbb{R}, \quad C_{\mu_A}(\xi_1, \eta_1) = C_\mu(A^\dagger \xi_1, A^\dagger \eta_1),$$

where $A^\dagger : V_1^* \to V_0^*$ is the adjoint of A defined by

$$\langle A^\dagger \xi_1, v_0\rangle = \langle \xi_1, Av_0\rangle, \; \forall v_0 \in V_0, \;\; \xi_1 \in V_1^*.$$

Indeed, let $\xi_1 \in V_1^*$, then $(\xi_1)_{\#}\mu_A = (\xi_1)_{\#}(A_{\#})\mu = (\xi_1 \circ A)_{\#}\mu$ and observe that

$$\xi_1 \circ A = A^\dagger \xi_1 \in V_0^*.$$

Hence,

$$\mathbb{P}_{\xi_1} = \boldsymbol{\Gamma}_{m[A^*\xi_1], C_\mu(A^\dagger \xi_1, A^\dagger \xi_1)}[dx]$$

Remark 1.1.18. Suppose that $\boldsymbol{U}_0, \boldsymbol{U}_1$ are Euclidean spaces and $\mu \in \mathscr{G}(\boldsymbol{U}_0)$. If $A : \boldsymbol{U}_0 \to \boldsymbol{U}_1$ is a linear operator then the variance operator of $\mu_A = A_{\#}\mu$ is

$$\operatorname{Var}_{\mu_A} = A \operatorname{Var}_\mu A^* : \boldsymbol{U}_1 \to \boldsymbol{U}_1. \tag{1.1.10}$$

In particular, if $C : \boldsymbol{U}_0 \to \boldsymbol{U}_0$ is a symmetric, nonnegative operator and $C^{1/2}$ denotes its nonnegative square root, then the probability measure $\Gamma_C := (C^{1/2})_{\#}\Gamma_{\boldsymbol{U}_0}$ is Gaussian, centered and its variance is C. We deduce that for any $\boldsymbol{u}_0 \in \boldsymbol{U}$ the pushforward of $\Gamma_{\boldsymbol{U}_0}$ via the affine map $\mathcal{T}_{\boldsymbol{u}_0} C^{1/2}$ is a Gaussian measure with variance C and mean $\boldsymbol{u}_0$. Thus, for any symmetric nonnegative operator C on $\boldsymbol{U}_0$, and any $m \in \boldsymbol{U}_0$, there exists a unique Gaussian measure μ on $\boldsymbol{U}_0$ with mean m and variance C. We denote it by $\Gamma_{m,C}$. More precisely, $\gamma_{m,C} = (\mathcal{T}_m \circ \sqrt{C})_{\#}\Gamma_{\boldsymbol{U}_0}$. □

Definition 1.1.19. Let V be a finite dimensional vector space and μ a Gaussian measure on V. We say that μ is nondegenerate if $\mu\big[\,\mathcal{O}\,\big] \neq 0$, for any open subset $\mathcal{O} \subset V$. □

Proposition 1.1.20. *Let V be a finite dimensional vector space and $\mu \in \mathcal{G}(V)$ the following are equivalent.*

(i) *The measure μ is nondegenerate.*

(ii) *The covariance form* C_μ *is nondegenerate , i.e.,*

$$C_\mu(\xi,\eta)=0,\ \ \forall\eta\in V^*\Longleftrightarrow\xi=0.$$

Proof. Clearly it suffices to consider only centered Gaussian measures. Fix an inner product on V so we can identify C_μ with a symmetric operator $C:V\to V$. Set $n=\dim V$ and fix an orthonormal basis of V that diagonalizes C,

$$C=\operatorname{Diag}(\lambda_1,\dots,\lambda_n).$$

Using this basis we identify V isometrically with $\mathbb{R}^n$. We have

$$\mu=(C^{1/2})_\#\Gamma_{\mathbb{1}_n}=(C^{1/2})_\#\big(\boldsymbol{\Gamma}_1\otimes\cdots\otimes\boldsymbol{\Gamma}_1\big)$$

$$=\big((\mathcal{R}_{\sqrt{\lambda_1}})_\#\boldsymbol{\Gamma}_1\big)\otimes\cdots\otimes\big((\mathcal{R}_{\sqrt{\lambda_n}})_\#\boldsymbol{\Gamma}_1\big)\stackrel{(1.1.1)}{=}\boldsymbol{\Gamma}_{\lambda_1}\otimes\cdots\boldsymbol{\Gamma}_{\lambda_n}.$$

We see that

$$\mu\ \text{ is nondegenerate}\Longleftrightarrow\prod_i\lambda_i\neq0\Longleftrightarrow C\ \text{ is invertibe.}$$

□

Remark 1.1.21. Suppose that μ is centered Gaussian measure on the Euclidean space $\boldsymbol{U}$ with inner product $(-,-)$. Denote by C the variance operator of μ, $C\in\mathbf{Sym}(\boldsymbol{U})$. The proof of Proposition 1.1.20 shows that the measure μ is supported on $(\ker C)^\perp$, i.e., $\mu\big[\,\mathcal{O}\,\big]=0$, for any open subset in $\boldsymbol{U}\setminus(\ker C)^\perp$.

The argument used in the proof of Proposition 1.1.20 shows that if μ is nondegenerate and thus C is invertible, then μ is absolutely continuous with respect to the Lebesgue measure $\boldsymbol{\lambda}_{\boldsymbol{U}}$ and

$$\mu\big[\,d\boldsymbol{u}\,\big]=\frac{1}{\sqrt{\det(2\pi C)}}e^{-\frac{1}{2}(C^{-1}\boldsymbol{u},\boldsymbol{u})}\boldsymbol{\lambda}_{\boldsymbol{U}}\big[\,d\boldsymbol{u}\,\big].$$ □

Proposition 1.1.22. *Let* $\boldsymbol{U}$ *be a finite dimensional Euclidean vector space with inner product* $(-,-)$*. Suppose that* $(\mu_n)_{n\in\mathbb{N}}$ *a sequence of Gaussian measures on* $\boldsymbol{U}$*. We set* $m_n:=m[\mu_n]$, $C_n:=\operatorname{Var}_{\mu_n}$*. The following are equivalent.*

(i) *The sequence* (μ_n) *converges weakly to a probability measure* μ_∞*.*

(ii) *The sequences* (m_n) *and* (C_n) *converge.*

(iii) *The sequence* (μ_n) *converges weakly to a Gaussian measure* μ_∞*.*

Proof. Clearly (iii) ⇒ (i). Note that if $m_n\to m$ and $C_n\to C$, then $\mu_n=\boldsymbol{\Gamma}_{m_n,C_n}$

$$\widehat{\boldsymbol{\Gamma}}_{m_n,C_n}(\xi)\to\widehat{\boldsymbol{\Gamma}}_{m,C}(\xi),\ \ \forall\xi\in\boldsymbol{U}^*.$$

Lévy's theorem implies that the sequence (μ_n) converges weakly to $\boldsymbol{\Gamma}_{m,C}$. Thus (ii) ⇒ (iii) so it suffices to prove (i) ⇒ (ii).

Condition (i) implies that

$$\widehat{\mu}_n(\xi)\to\widehat{\mu}(\xi),\ \ \forall\xi\in V^*.$$

For any $\xi\in\boldsymbol{U}^*$ the Fourier transform of the measure $\xi_\#\mu_n$ is

$$\widehat{\xi_\#\mu}_n(t)=\widehat{\mu}_n(t\xi)\to\widehat{\mu}(t\xi)=\widehat{\xi_\#\mu}(t\xi).$$

Proposition 1.1.2 implies that $\xi_{\#}\mu$ is Gaussian and

$$m_n[\xi_n] \to m[\xi], \quad v_n[\xi] \to v[\xi].$$

□

Definition 1.1.23. Let V be a finite dimensional real vector space.

(i) A random vector $Z : (\Omega, \mathcal{S}, \mathbb{P}) \to V$ is called *Gaussian* if its distribution $\mathbb{P}_Z$ is a Gaussian measure on V. The covariance form of Z, denoted by Cov_Z is the covariance form of its distribution.

(ii) The Gaussian vector Z is called *centered*, respectively *nondegenerate* if its distribution is such.

(iii) The real valued random variables $(X_1, \dots, X_n)$ are said to be *jointly Gaussian* if the random vector $(X_1, \dots, X_n)$ is Gaussian.

□

Remark 1.1.24. (a) The Cov_Z is a bilinear form on V^*. Identifying $(V^*)^*$ with V in the canonical fashion we can view Cov_Z as an element of $V \otimes V$. As such, it can be given the more compact geometric description

$$\mathrm{Cov}_Z = \mathbb{E}\big[\, Z \otimes Z \,\big] - \mathbb{E}\big[\, Z \,\big] \otimes \mathbb{E}\big[\, Z \,\big].$$

(b) If V is equipped with an inner product, then we can identify Cov_Z with a symmetric, nonnegative operator $\mathrm{Var}\big[\, X \,\big]$ uniquely determined by the equalities

$$\big(\, v_1, \mathrm{Var}\big[\, X \,\big] v_2 \,\big) = \mathrm{Cov}\big[\, v_1^{\dagger}(X), v_2^{\dagger}(X) \,\big], \quad \forall v_1, v_2 \in V.$$

Moreover, if $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ is an *orthonormal* basis of V, then we can write

$$Z = \sum_{i=1}^{n} Z_i \boldsymbol{e}_i, \quad Z_i \in L^2\big(\, \Omega, \mathcal{S}, \mathbb{P} \,\big)$$

and the variance operator of Z is described by the Gramian matrix of the Gaussian random variables $\widehat{Z}_i = Z_i - \mathbb{E}\big[\, Z_i \,\big]$, $i = 1, \dots, n$. This is the $n \times n$ matrix

$$G\big(\, \widehat{Z}_1, \dots, \widehat{Z}_n \,\big) = \Big(\, \mathbb{E}\big[\, \widehat{Z}_i \widehat{Z}_j \,\big] \,\Big)_{1 \le i,j \le n}.$$

We see that Z is nondegenerate if and only if the random variables $\widehat{Z}_i$ are linearly independent.

□

Suppose that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are finite dimensional vector spaces. Given random vectors

$$X : (\Omega, \mathcal{S}, \mathbb{P}) \to \boldsymbol{X}, \quad Y : (\Omega, \mathcal{S}, \mathbb{P}) \to \boldsymbol{Y}$$

we define the *covariance form* of Y and X to be the bilinear form

$$C_{Y,X} : \boldsymbol{Y}^* \times \boldsymbol{X}^* \to \mathbb{R}$$

given by

$$C_{Y,X}(\eta, \xi) = \mathrm{Cov}\big[\, \langle \eta, Y \rangle, \langle \xi, X \rangle \,\big], \quad \forall \eta \in \boldsymbol{Y}^*, \ \xi \in \boldsymbol{X}^*.$$

If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are equipped with inner products $(-,-)_{\boldsymbol{X}}$ and respectively $(-,-)_{\boldsymbol{Y}}$, then we can identify $\operatorname{Cov}\big[\,Y,X\,\big]$ with a linear operator $\operatorname{Cov}[Y,X]:\boldsymbol{X}\to\boldsymbol{Y}$ uniquely determined by the condition

$$\big(\,y,\operatorname{Cov}[Y,X]x\,\big)_Y=\operatorname{Cov}\big[\,(y,Y)_{\boldsymbol{Y}},\ (x,X)_{\boldsymbol{X}}\,\big],\ \ \forall x\in\boldsymbol{X},\ \ y\in\boldsymbol{Y}.$$

The operator $\operatorname{Cov}[Y,X]$ is called the *covariance operator* of Y and X.

Concretely, if $(\boldsymbol{e}_i)_{i\in I}$ and $(\boldsymbol{f}_j)_{j\in J}$ are *orthonormal* bases of $\boldsymbol{X}$ and respectively $\boldsymbol{Y}$, and we set $X_i:=(\boldsymbol{e}_i,X)_{\boldsymbol{X}}$, $Y_j:=(\boldsymbol{f}_j,Y)_{\boldsymbol{Y}}$, then in these bases the operator $\operatorname{Cov}[Y,X]$ is described by matrix $(c_{ji})_{(j,i)\in J\times I}$, where $c_{ji}:=\operatorname{Cov}[Y_j,X_i]$. Hence

$$\operatorname{Cov}[Y,X]\boldsymbol{e}_i=\sum_j c_{ji}\boldsymbol{f}_j.$$

Let us observe that $\operatorname{Cov}[X,X]=\operatorname{Var}\big[\,X\,\big]$ and that $\operatorname{Cov}[X,Y]:\boldsymbol{Y}\to\boldsymbol{X}$ is the adjoint of $\operatorname{Cov}[Y,X]$

$$\operatorname{Cov}[X,Y]=\operatorname{Cov}[Y,X]^*.$$

Note that if $T:\boldsymbol{X}\to\boldsymbol{U}$ is a linear map between Euclidean spaces, then

$$\operatorname{Cov}[Y,TX]=\operatorname{Cov}[Y,X]\circ T^*:\boldsymbol{U}\to\boldsymbol{Y}.$$

The random vectors X,Y are said to be *jointly Gaussian* if the random vector

$$X\oplus Y:(\Omega,\mathcal{S},\mathbb{P})\to\boldsymbol{X}\oplus\boldsymbol{Y}$$

is Gaussian. If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are equipped with inner products, then $\boldsymbol{X}\oplus\boldsymbol{Y}$ is equipped the the direct sum of these inner products and in this case $\operatorname{Var}\big[\,X\oplus Y\,\big]:\boldsymbol{X}\oplus\boldsymbol{Y}\to\boldsymbol{X}\oplus\boldsymbol{Y}$ admits the bloc decomposition

$$\operatorname{Var}\big[\,X\oplus Y\,\big]=\left[\begin{array}{cc}\operatorname{Var}\big[\,X\,\big] & \operatorname{Cov}[X,Y]\\ \operatorname{Cov}[Y,X] & \operatorname{Var}\big[\,Y\,\big]\end{array}\right].$$

We deduce from the above the following very convenient fact.

Proposition 1.1.25. *Suppose that the random vectors X,Y are* jointly Gaussian*. Then X,Y are independent iff the covariance operator* $\operatorname{Cov}[Y,X]$ *is trivial.* □

Suppose that $\boldsymbol{W}$ is an m-dimensional real Euclidean space with inner product $(-,-)$. Denote by $S_1(\boldsymbol{W})$ the unit sphere in $\boldsymbol{V}$ and by $\mathbf{Sym}(\boldsymbol{W})$ the space of symmetric operators $\boldsymbol{V}\to\boldsymbol{V}$ and by $\mathbf{Sym}_{\geq 0}(\boldsymbol{V})$ the cone of nonnegative ones. For $A\in\mathbf{Sym}_{\geq 0}(\boldsymbol{W})$ we denote by $\boldsymbol{\Gamma}_A$ the centered Gaussian measure on $\boldsymbol{W}$ with variance A.

The space $\mathbf{Sym}(\boldsymbol{W})$ is equipped with an operator norm $\|-\|_{\mathrm{op}}$

$$\|A\|_{\mathrm{op}}:=\sup_{\|\boldsymbol{w}\|=1}\|A\boldsymbol{w}\|=\ \sup\operatorname{Spec}\big(\,|A|\,\big).$$

There is another trace norm

$$\|A\|_1:=\operatorname{tr}(|A|).$$

Note that for any $A\in\mathbf{Sym}(\boldsymbol{W})$ we have

$$\|A\|_{\mathrm{op}}\leq\|A\|_1\leq m\|A_{\mathrm{op}}\|.$$

We have a natural map $\mathbf{Sym}_{\geq 0}(\boldsymbol{W})\to\mathbf{Sym}_{\geq 0}(\boldsymbol{W})$, $A\mapsto A^{1/2}$. We will need the following result, [70, Prop.2.1].

Proposition 1.1.26. *For any $\mu > 0$ and $\forall A, B \in \mathbf{Sym}_{\geq 0}(\boldsymbol{W})$, such that $A^{1/2} + B^{1/2} \geq \mu \mathbb{1}$*

$$\mu \| A^{1/2} - B^{1/2} \|_1 \leq \| A - B \|_1^{1/2}, \tag{1.1.11}$$

□

Any continuous function $f : \boldsymbol{W} \to \mathbb{R}$ with at most polynomial growth defines a map

$$\mathbf{Sym}_{\geq 0}(\boldsymbol{W}) \ni A \mapsto \mathcal{I}_A(f) := \int_{\boldsymbol{W}} f(\boldsymbol{w}) \boldsymbol{\Gamma}_A \big[d\boldsymbol{w} \big] \in \mathbb{R}.$$

Lemma 1.1.27. *Fix $\mu_0 > 0$ and suppose that $f : \boldsymbol{V} \to \mathbb{R}$ is a locally Lipschitz function that is homogeneous of degree $k \geq 1$. Denote by $\operatorname{Lip}(f)$ the Lipschitz constant of the restriction of f to the unit ball, i.e.,*

$$\operatorname{Lip}(f) := \sup_{\substack{\|u\|, \|\boldsymbol{v}\| \leq 1 \\ \boldsymbol{u} \neq \boldsymbol{v}}} \frac{|f(\boldsymbol{u}) - f(\boldsymbol{v})|}{\|\boldsymbol{u} - \boldsymbol{v}\|}$$

Then there exists a constant $C = C(m, k) > 0$ with the following property for and $R \geq 0$ and any $A, B \in \mathbf{Sym}_{\geq 0}(\boldsymbol{V})$ such that

$$A^{1/2} + B^{1/2} \geq \mu_0 \mathbb{1}, \;\; \|A^{1/2}\|_{\mathrm{op}}, \; \|B^{1/2}\|_{\mathrm{op}} \leq R$$

$$\big| \mathcal{I}_A(f) - \mathcal{I}_B(f) \big| \leq \frac{\operatorname{Lip}(f) R^k C(m, k)}{\mu_0} \|A - B\|_{\mathrm{op}}^{1/2}. \tag{1.1.12}$$

In other words, $A \mapsto \mathcal{I}_A(f)$ is locally Hölder continuous with exponent $1/2$ in the open set $\mathbf{Sym}_{>0}\big(\boldsymbol{V}\big)$.

Proof. If we denote by $B_R(\boldsymbol{V})$ the closed ball of radius R, then the homogeneity of f implies that

$$\big[f(\boldsymbol{u}) - f(\boldsymbol{v}) \big| \leq \operatorname{Lip}(f) R^k \|\boldsymbol{u} - \boldsymbol{v}\|, \;\; \forall \boldsymbol{u}, \boldsymbol{v} \in B_R(\boldsymbol{W}). \tag{1.1.13}$$

Note that

$$\mathcal{I}_A(f) = \int_{\boldsymbol{W}} f\big(A^{1/2} \boldsymbol{w} \big) \Gamma_{\mathbb{1}} \big[d\boldsymbol{w} \big],$$

so

$$\big| \mathcal{I}_A(f) - \mathcal{I}_B(f) \big| \leq \int_{\boldsymbol{W}} \big| f\big(A^{1/2}\boldsymbol{w} \big) - f\big(B^{1/2}\boldsymbol{w} \big) \big| \boldsymbol{\Gamma}_{\mathbb{1}} \big[d\boldsymbol{w} \big]$$

$$= \underbrace{\frac{1}{(2\pi)^{m/2}} \left(\int_0^\infty r^{m+k-1} e^{-r^2/2} dr \right)}_{C_{m,k}} \int_{S_1(\boldsymbol{V})} \big| f\big(A^{1/2}\boldsymbol{w} \big) - f\big(B^{1/2}\boldsymbol{w} \big) \big| \operatorname{vol}_{S_1(\boldsymbol{W})} \big[d\boldsymbol{w} \big]$$

$(\|A^{1/2}\boldsymbol{w}\|, \;\; \|B^{1/2}\boldsymbol{w}\| \leq R)$

$$\stackrel{(1.1.13)}{\leq} C_{m,k} \operatorname{Lip}(f) R^k \int_{S_1(\boldsymbol{W})} \|A^{1/2} - B^{1/2}\|_{\mathrm{op}} \operatorname{vol}_{S^1(\boldsymbol{V})} \big[d\boldsymbol{w} \big]$$

$$\leq C_{m,k} \operatorname{Lip}(f) R^k \operatorname{vol} \big[S_1(\boldsymbol{W}) \big] \int_{S_1(\boldsymbol{W})} \|A^{1/2} - B^{1/2}\|_1 \operatorname{vol}_{S^1(\boldsymbol{V})} \big[d\boldsymbol{w} \big]$$

$$\stackrel{(1.1.11)}{\leq} \frac{Z(m,k) \operatorname{Lip}(f) R^k}{\mu_0} \|A - B\|_1^{1/2} \leq \frac{Z(m,k) m^{1/2} \operatorname{Lip}(f) R^k}{\mu_0} \|A - B\|_{\mathrm{op}}^{1/2}.$$

□

Remark 1.1.28. Observe that

$$\operatorname{tr} A \leq R^2 \implies \|A\|_{\mathrm{op}} \leq R.$$

Set

$$a := \inf \operatorname{Spec}(A), \quad b := \inf \operatorname{Spec}(B).$$

We have

$$\begin{aligned}\inf \operatorname{Spec}(A^{1/2}+B^{1/2}) &= \inf_{\|\boldsymbol{u}\|=1} \big((A^{1/2}+B^{1/2})\boldsymbol{u}, \boldsymbol{u} \big)\\ &\geq \inf \operatorname{Spec}(A^{1/2}) + \inf \operatorname{Spec}(B^{1/2}) = \sqrt{a}+\sqrt{b} \geq \sqrt{a+b}.\end{aligned}$$

Hence

$$a+b \geq \mu_0^2 \implies A^{1/2}+B^{1/2} \geq \mu_0 \mathbb{1}.$$

□

Lemma 1.1.29. *Suppose that $f : \boldsymbol{W} \to \mathbb{R}$ is a continuous function that is homogeneous of degree $k \geq 1$. Set*

$$M(f) := \sup_{\|\boldsymbol{w}\| \leq 1} |f(\boldsymbol{w})|.$$

Then there exists $C = C(m,k) > 0$ such that $\forall A \in \mathbf{Sym}_{\geq 0}(\boldsymbol{V})$

$$\big| \mathcal{I}_A(f) \big| \leq \mathcal{I}_A(|f|) \leq C(m,k) M(f) \|A\|_{\mathrm{op}}^{k/2}. \tag{1.1.14}$$

Proof. Note that

$$\sup_{\|\boldsymbol{w}\| \leq R} |f(\boldsymbol{w})| = M(f) R^k.$$

As in the proof of Lemma 1.1.27 we have

$$\begin{aligned}\mathcal{I}_A(|f|) &= \int_{\boldsymbol{W}} f\big(A^{1/2}\boldsymbol{w} \big) \boldsymbol{\Gamma}_{\mathbb{1}}\big[d\boldsymbol{w} \big]\\ &= \underbrace{\frac{1}{(2\pi)^{m/2}} \left(\int_0^\infty r^{m+k-1} e^{-r^2/2} dr \right)}_{=:C_{m,k}} \int_{S_1(\boldsymbol{W})} \big| f\big(A^{1/2}\boldsymbol{w} \big) \big| \operatorname{vol}_{S_1(\boldsymbol{V})} \big[d\boldsymbol{w} \big]\end{aligned}$$

$(\|A^{1/2}\boldsymbol{w}\| \leq \|A^{1/2}\|_{\mathrm{op}} \|\boldsymbol{w}\|)$

$$\leq C_{m,k} M(f) \|A^{1/2}\|_{\mathrm{op}}^k \operatorname{vol}\big[S_1(\boldsymbol{V}) \big] = C(m,k) M(f) \|A\|_{\mathrm{op}}^{k/2}.$$

□

Corollary 1.1.30. *Suppose that $f : \boldsymbol{W} \to \mathbb{R}$ is a continuous function that is homogeneous of degree $k \geq 1$. Suppose that $A, B \in \mathbf{Sym}_{\geq 0}(\boldsymbol{W})$ and $B \leq A$. Then*

$$\big| \mathcal{I}_B(f) \big| \leq \mathcal{I}_B(|f|) \leq C(m,k) M(f) \|B\|_{\mathrm{op}}^{k/2} \leq C(m,k) M(f) \|A\|_{\mathrm{op}}^{k/2}. \tag{1.1.15}$$

Proof. Indeed, $0 \leq B \leq A \implies \|B\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}}$. □

Proposition 1.1.31. *Let $X_n : (\Omega, \mathcal{S}, \mathbb{P}) \to \boldsymbol{U}$ be a sequence of Gaussian vectors valued in the m-dimensional Euclidean space $\boldsymbol{U}$. Assume that*

(i) *for any $m < n$ the vectors X_m, X_n are jointly Gaussian and,*

(ii) *the vectors X_n converge a.s. to a random vector $X : (\Omega, \mathcal{S}, \mathbb{P}) \to \boldsymbol{U}$.*

Then X is Gaussian and $X_n \to X$ in L^p, $\forall p \in [1, \infty)$.

Proof. The vectors X_n converge in distribution to X and Proposition 1.1.22 shows that

$$\mathbb{E}\big[\, X_n \,\big] \to \mathbb{E}\big[\, X \,\big], \;\; \operatorname{Var}\big[\, X_n \,\big] \to \operatorname{Var}\big[\, X \,\big].$$

Replacing X_n with $\bar{X}_n = X_n - \mathbb{E}\big[\, X_n \,\big]$ we can assume that Xcentered. Set $Y_n = X_n - X$. Note that the Gaussian vector $(X_n - X_m)$ converges a.s. to Y_n as $m \to \infty$ so Y_n is a Gaussian vector as well. Moreover $Y_n \to 0$ a.s.. Set $C_n := \operatorname{Var}\big[\, Y_n \,\big]$ and $K_n = \|C_n\|_{\mathrm{op}}$. Then $K_n \to 0$ as $n \to \infty$.

Let $p \in [1, \infty)$. Lemma 1.1.29 implies that

$$\mathbb{E}\big[\, |Y_n|^p \,\big] = \int_{\boldsymbol{U}} |\boldsymbol{u}|^p \boldsymbol{\Gamma}_{C_n}[d\boldsymbol{u}\,] \leq C(m,p) K_n^{p/2},$$

where $C(m,p) > 0$ depends only on $m = \dim \boldsymbol{U}$ and $p \geq 1$. This proves that $Y_n \to 0$ in L^p. □

1.1.2. Gaussian regression. Suppose that X, Y are two L^2- random vectors valued in the Euclidean spaces $\boldsymbol{X}$ and respectively $\boldsymbol{Y}$. Denote by $\mathbf{Aff}(\boldsymbol{X}, \boldsymbol{Y})$ the space of affine maps $\boldsymbol{X} \to \boldsymbol{Y}$. The classical *least square approximation* gives an explicit description of an affine map $\mathcal{A}_0 : \boldsymbol{X} \to \boldsymbol{Y}$ such that

$$\mathbb{E}\big[\, \|Y - \mathcal{A}_0 X\|^2 \,\big] \leq \mathbb{E}\big[\, \|Y - \mathcal{A} X\|^2 \,\big], \;\; \forall \mathcal{A} \in \mathbf{Aff}(\boldsymbol{C}, \boldsymbol{Y}).$$

The $\boldsymbol{Y}$-valued random vector $\mathcal{A}_0 X$ is called the *linear regression.*

If X, Y are two arbitrary random vectors, then the conditional expectation $\mathbb{E}\big[\, Y \,\|\, X \,\big]$ is some measurable function of X. The next result shows that *when X, Y are jointly Gaussian* and X is nondegenerate, there exists a unique affine map $\mathcal{A}_0$ such that $\mathcal{A}_0 X = \mathbb{E}\big[\, Y \,\|\, X \,\big]$. Moreover $\mathcal{A}_0$ is the solution of the above minimization problem.

Proposition 1.1.32 (Gaussian regression formula)**.** *Suppose that X, Y are Gaussian vectors valued in the Euclidean spaces $\boldsymbol{X}$ and respectively $\boldsymbol{Y}$. Denote by m_X and respecitvel m_Y the meand of X and respectively m_Y. Assume additionally that*

(i) *the random vectors X, Y are jointly Gaussian and,*
(ii) *X is nondegenerate.*

Define the regression operator

$$R_{Y,X} : \boldsymbol{X} \to \boldsymbol{Y}, \;\; R_{Y,X} := \operatorname{Cov}[Y, X] \operatorname{Var}[X]^{-1} \tag{1.1.16}$$

Then the following hold.

(a) The conditional expectation $\mathbb{E}\big[\, Y \,\|\, X \,\big]$ is a linear function of X described by the linear regression formula

$$\mathbb{E}\big[\, Y \,\|\, X \,\big] = m_Y - R_{Y,X} m_X + R_{Y,X} X. \tag{1.1.17}$$

(b) For any $x \in X$

$$\mathbb{E}\big[\, Y \big|\, X = x \,\big] = m_Y - R_{Y,X} m_X + R_{Y,X} x.$$

(c) The random vector vector $Z = Y - \mathbb{E}\big[\, Y \,\|\, X \,\big]$ is Gaussian and independent of X. It has mean 0 and variance operator

$$\Delta_{Y,X} = \operatorname{Var}\big[\, Y \,\big] - D_{Y,X} : \boldsymbol{Y} \to \boldsymbol{Y}, \;\; D_{Y,X} = \operatorname{Cov}[Y, X] \operatorname{Var}[X]^{-1} \operatorname{Cov}[X, Y]. \tag{1.1.18}$$

Moreover, for any bounded measurable function $f : \boldsymbol{Y} \to \mathbb{R}$ and any $x \in \boldsymbol{X}$ we have

$$\mathbb{E}\big[\, f(Y) \big|\, X = x \,\big] = \mathbb{E}\big[\, f\big(\, Z + m_Y - R_{Y,X} m_X + R_{Y,X} x \,\big)\,\big]. \tag{1.1.19}$$

In particular, if X and Y are centered we have

$$\mathbb{E}\big[\, f(Y) \big|\, X = x \,\big] = \mathbb{E}\big[\, f\big(\, Z + R_{Y,X} x \,\big)\,\big]. \tag{1.1.20}$$

Proof. Assume first that both X and Y are centered. Set

$$Z =: Y - R_{Y,X} X,$$

where $R_{Y,X}$ is defined in (1.1.16). Assumption (i) implies that Z is also a centered Gaussian vector.

Let $(\boldsymbol{e}_i)_{i \in I}$ and $(\boldsymbol{f}_\alpha)_{\alpha \in A}$ are orthonormal bases of $\boldsymbol{X}$ and respectively $\boldsymbol{Y}$. Set

$$X_i := (\boldsymbol{e}_i, X)_{\boldsymbol{X}},\;\; Y_\alpha := (\boldsymbol{f}_\alpha, Y)_{\boldsymbol{Y}},\;\; Z_\alpha := (\boldsymbol{f}_\alpha, Z)_{\boldsymbol{Y}},$$

and

$$V(X)_{ij} := \mathbb{E}\big[\, X_i X_j \,\big],\;\; C_{\alpha i} := \mathbb{E}\big[\, Y_\alpha X_i \,\big] = C_{i\alpha},\;\; V(Y)_{\alpha\beta} := \mathbb{E}\big[\, Y_\alpha Y_\beta \,\big].$$

The matrix $\big(\, V(X)_{ij} \,\big)_{i,j \in I}$ describes the variance operator of X, the matrix $\big(\, V(Y)_{\alpha\beta} \,\big)_{\alpha,\beta \in A}$ describes the variance operator of Y and the matrix $\big(\, C_{\alpha i} \,\big)_{\alpha \in A, i \in I}$ defines the covariance operator $\mathrm{Cov}[Y, X]$. We denote by $V(X)^{-1}_{ij}$ the entries of $\mathrm{Var}\,\big[\, X \,\big]^{-1}$ and by $D_{\alpha\beta}$ the entries of $D_{Y,X} = \mathrm{Cov}[Y, X]\,\mathrm{Var}[X]^{-1}\,\mathrm{Cov}[X, Y]$. We have

$$R_{X,Y} X = \sum_\alpha \Big(\, \sum_i R_{\alpha i} X_i \,\Big) X_i \,\Big) \boldsymbol{f}_\alpha,$$

where

$$R_{\alpha i} = \sum_j C_{\alpha j} V(X)^{-1}_{ji}.$$

Hence

$$Z_\alpha = Y_\alpha - \sum_i R_{\alpha i} X_i,\;\; Z_\beta = Y_\beta - \sum_j R_{\beta j} X_j,$$

$$\mathbb{E}\big[\, Z_\alpha Z_\beta \,\big] = V(Y)_{\alpha\beta} - \sum_j R_{\beta j} C_{\alpha j} - \sum_i R_{\alpha i} C_{i\beta} + \sum_{i,j} R_{\alpha i} V_{ij} R_{\beta j}.$$

We have

$$\sum_i \sum_j R_{\alpha i} V_{ij} R_{\beta j} = \sum_i \sum_j \Big(\, \sum_k C_{\alpha k} V(X)^{-1}_{ki} V_{ij} \,\Big) R_{\beta_j}$$

$$= \sum_j \Big(\, \sum_k C_{\alpha k} \delta_{kj} \,\Big) R_{\beta j} = \sum_k C_{\alpha k} R_{\beta_k} = \sum_k R_{\beta_k} C_{k\alpha} = D_{\beta\alpha} = D_{\alpha\beta}.$$

A similar but simpler computation shows that

$$\sum_j R_{\beta j} C_{\alpha j} = D_{\beta\alpha} = D_{\alpha\beta} = \sum_i R_{\alpha i} C_{i\beta}.$$

Thus $\Delta_{Y,X} = \mathrm{Var}\,\big[\, Y \,\big] - D_{Y,X}$ is the covariance operator of Z.

An elementary computation shows that.

$$\mathbb{E}\big[\, Z_\alpha X_i \,\big] = 0,\;\; \forall \alpha, i$$

and assumption (i) implies that X and Z are independent centered Gaussian vectors. Clearly Z is an X-measurable random vector. If S is and X-measurable event, then

$$\mathbb{E}\big[\, Z\boldsymbol{I}_F \,\big] = \mathbb{E}\big[\, Z \,\big]\mathbb{P}\big[\, F \,\big] = 0.$$

Hence

$$\mathbb{E}\big[\, Y\boldsymbol{I}_F \,\big] - \mathbb{E}\big[\, R_{Y,XS}\boldsymbol{I}_F \,\big] = \mathbb{E}\big[\, Z\boldsymbol{I}_F \,\big] = 0$$

so that

$$R_{Y,X}X = \mathbb{E}\big[\, Y \,\|\, X \,\big]$$

and

$$\mathbb{E}\big[\, Y \big|\, X = x \,\big] = R_{Y,X}x.$$

Now let $f : \boldsymbol{Y} \to \mathbb{R}$ be a bounded measurable function. Then $Y = \mathbb{E}\big[\, Y \,\|\, X \,\big] + Z$, with $\mathbb{E}\big[\, Y \,\|\, X \,\big], Z$ independent. Then

$$\begin{aligned}\mathbb{E}\big[\, f(Y) \big|\, X = x \,\big] &= \mathbb{E}\big[\, f\big(\, Z + \mathbb{E}[Y \,\|\, X]\,\big) \big|\, X = x \,\big]\\ &= \mathbb{E}\big[\, f(Z + \mathbb{E}[Y \big|\, X = x])\,\big] = \mathbb{E}\big[\, f(Z + R_{Y,X}x)\,\big].\end{aligned}$$

This proves the Proposition 1.1.32 when both X and Y are centered.

We now reduce the general case to the centered case. Consider the centered vectors

$$\bar{X} := X - m_X,\;\; \bar{Y} = Y - m_Y.$$

Then

$$\begin{aligned}R_{Y,X} &= R_{\bar{Y},\bar{X}},\\ \mathbb{E}\big[\, Y \,\|\, X \,\big] &= m_Y + \mathbb{E}\big[\, \bar{Y} \,\|\, X \,\big] = m_Y + \mathbb{E}\big[\, \bar{Y} \,\|\, \bar{X} \,\big]\\ &= m_Y + R_{Y,X}\bar{X} = m_Y - R_{Y,X}m_X + R_{Y,X}X.\end{aligned}$$

If we set

$$\bar{Z} = \bar{Y} - R_{Y,X}\bar{X} = Y - m_Y + R_{Y,X}m_X - R_{Y,X}X = Y - \mathbb{E}\big[\, Y \,\|\, X \,\big],$$

then $\bar{Z}$ is independent of $\bar{X}$ and thus also of X. □

Remark 1.1.33. (a) Let $\boldsymbol{X}$, $\boldsymbol{Y}$, X and Y be as in the above proposition. Assume additionally that X and Y are centered. Sometimes we will use the notation

$$\mathrm{Var}\big[\, Y | X = 0 \,\big] := \Delta_{Y,X}.$$

Note that

$$\mathrm{Var}\big[\, Y | X = 0 \,\big] = \mathrm{Var}\big[\, Y \,\big] - \mathrm{Cov}[Y,X]\,\mathrm{Var}\big[\, X \,\big]^{-1}\mathrm{Cov}[X,Y\,] \leq \mathrm{Var}\big[\, Y \,\big], \tag{1.1.21}$$

since the symmetric operator $\mathrm{Cov}[Y,X]\,\mathrm{Var}\big[\, X \,\big]^{-1}\mathrm{Cov}[X,Y\,]$ is nonnegative.

(b)Suppose that $\boldsymbol{U}$ is another Euclidean space and $T : \boldsymbol{X} \to \boldsymbol{U}$ is a linear isomorphism. Then for any positively homogeneous measurable function $f : \boldsymbol{Y} \to \mathbb{R}$ we have

$$\mathbb{E}\big[\, f(Y) \big|\, X = 0 \,\big] = \mathbb{E}\big[\, f(Y) \big|\, TX = 0 \,\big].$$

To see this is suffices to show that $\Delta_{Y,X} = \Delta_{Y,TX}$ This happens iff

$$\mathrm{Cov}[Y,X]\,\mathrm{Var}\big[\, X \,\big]^{-1}\mathrm{Cov}[X,Y] = \mathrm{Cov}[Y,TX]\,\mathrm{Var}\big[\, X \,\big]^{-1}\mathrm{Cov}[TX,Y].$$

Indeed,

$$\begin{aligned}\mathrm{Cov}[Y,TX] = \mathrm{Cov}[Y,X]T^*,\;\; \mathrm{Cov}[TX,Y] = T\,\mathrm{Cov}[X,Y],\\ \mathrm{Var}\big[\, TX \,\big] = T\,\mathrm{Var}\big[\, X \,\big]T^*.\end{aligned}$$

□

Proposition 1.1.34. *Suppose that $\boldsymbol{V}, \boldsymbol{U}$ are finite dimensional Euclidean spaces, V is a centered, $\boldsymbol{V}$-valued Gaussian vector, and $S : \boldsymbol{V} \to \boldsymbol{U}$ a linear surjection. Assume that the $\boldsymbol{U}$-valued Gaussian vector $S(V)$ is nondegenerate. Define $\boldsymbol{Y} = \ker S$, $\boldsymbol{X} = \boldsymbol{Y}^{\perp}$. Set*

$$L = (SS^*)^{-1/2} S : \boldsymbol{V} \to \boldsymbol{U}.$$

Denote by X and respectively Y the components of V along $\boldsymbol{X}$ and respectively $\boldsymbol{Y}$ so that $V = X + Y$ and $LV = LX$. Then the following hold

(i) *The Gaussian vectors LV and X are nondegenerate.*

(ii) *The Gaussian vectors $Y - \mathbb{E}\big[\, Y \,\|\, X \,\big]$, $V - \mathbb{E}\big[\, V \,\|\, LV \,\big]$ and $Y - \mathbb{E}\big[\, Y \,\|\, LV \,\big]$ have the same distribution and their common variance operator is $\Delta_{Y,X} : \boldsymbol{Y} \to \boldsymbol{Y}$ described in (1.1.18) . They are nondegenerate if and only if V is nondegenerate. Denote by $\Gamma_{\Delta_{Y,X}}$ the regression Gaussian measure, i.e., the centered Gaussian measure on $\boldsymbol{Y}$ with variance operator $\Delta_{Y,X}$.*

(iii) *If $f : \boldsymbol{V} \to \mathbb{R}$ is integrable with respect to the distribution of V, then*

$$\mathbb{E}\big[\, f(V) \,\big|\, L(V) = 0 \,\big] = \int_{\boldsymbol{Y}} f(y) \Gamma_{\Delta_{Y,X}}\big[\, dy \,\big] = \mathbb{E}\big[\, f(Y) \big|\, X = 0 \,\big]. \tag{1.1.22}$$

In particular, if the Gaussian vector V is nondegenerate and $f : \boldsymbol{V} \to (0, \infty)$ is a nonnegative, continuous homogeneous function whose restriction to $\ker S = \boldsymbol{Y}$ is nonzero, then

$$\mathbb{E}\big[\, f(V) \big|\, L(V) = 0 \,\big] = \int_{\ker S} f(y) \Gamma_{\Delta_{Y,X}}\big[\, dy \,\big] > 0. \tag{1.1.23}$$

Proof. The map $S\,|_{\boldsymbol{X}} : \boldsymbol{X} \to \boldsymbol{U}$ is an isomorphism and $S\,|_{\boldsymbol{X}}\,(X) = \boldsymbol{U} = S(V)$. Denote by P the orthogonal projection onto $\boldsymbol{X}$. Then $X = P(V)$, $Y = V - X$ and

$$S(V) = S(PV) = S(X).$$

Note that $S^*(\boldsymbol{U}) = \boldsymbol{X}$. Set $B := SS^* : \boldsymbol{U} \to \boldsymbol{U}$. The operator B is symmetric and positive definite. Observe that $L := B^{-1/2} S$.

Lemma 1.1.35. *The operator of L^* induces an isometry $\boldsymbol{U} \hookrightarrow \boldsymbol{V}$ with image*

$$L^*(\boldsymbol{U}) = (\ker L)^{\perp} = (\ker S)^{\perp} = \boldsymbol{X}.$$

Moreover $LL^* = \mathbb{1}_{\boldsymbol{U}}$.

Proof. Let $u_1, u_2 \in \boldsymbol{U}$. We have

$$(L^* u_1, L^* u_2) = (S^* B^{-1/2} u_1, S^* B^{-1/2} u_2)$$

$$= (SS^* B^{-1/2} u_1, B^{-1/2} u_2) = (B^{1/2} u_1, B^{-1/2} u_2) = (u_1, u_2).$$

Note that $LL^* = B^{-1/2} LL^* B^{-1/2} = \mathbb{1}$. □

If A denotes the variance operator of X, then the variance operator of $L(V) = L(X)$ is LAL^*. Moreover, $\mathrm{Cov}[Y, L(X)] = \mathrm{Cov}[Y, X] L^*$.

Denote by Q the variance operator of V. With respect to the decomposition $\boldsymbol{V} = \boldsymbol{X} \oplus \boldsymbol{Y}$ Q has the block form

$$Q = \begin{bmatrix} A & C^* \\ C & B \end{bmatrix}, \quad C = \mathrm{Cov}[Y, X], \quad B = \mathrm{Var}\big[\, Y \,\big].$$

Since X is nondegenerate, the operator A is invertible. Form the operator

$$\Delta_{Y,X} := \mathrm{Var}\big[\, Y \,\big] - \mathrm{Cov}[Y, X]\,\mathrm{Var}\big[\, X \,\big]^{-1}\mathrm{Cov}[Y, X] = B - CA^{-1}C^*$$

Then *Schur's complement formula* (see [**75**, Sec.0.8.5] or [**143**, Prop. 3.9])

$$\begin{bmatrix} \mathbb{1} & 0 \\ -CA^{-1} & \mathbb{1} \end{bmatrix} \cdot \begin{bmatrix} A & C^* \\ C & B \end{bmatrix} \cdot \begin{bmatrix} \mathbb{1} & -A^{-1}C^* \\ 0 & \mathbb{1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ & B - CA^{-1}C^* \end{bmatrix}$$

shows that $\det Q = \det A \cdot \det \Delta_{Y,X}$, so that $\det \Delta_{Y,X} \neq 0$ if and only if $\det Q \neq 0$, i.e., V is nondegenerate. Similarly

$$\begin{aligned} \Delta_{Y,LX} &= \mathrm{Var}\big[\, Y \,\big] - \mathrm{Cov}[Y, LX]\,\mathrm{Var}\big[\, LX \,\big]^{-1}\mathrm{Cov}[LX, Y] \\ &= B - CL^*(LAL^*\big)^{-1}LC^* = B - CA^{-1}C^* = \Delta_{Y,X}. \end{aligned}$$

since $LL^* = \mathbb{1}_{\boldsymbol{U}}$. This proves (ii).

From the equality

$$\mathbb{E}\big[\, V \,\|\, X \,\big] = \mathbb{E}\big[\, X + Y \,\|\, X \,\big] = \mathbb{E}\big[\, Y \,\|\, X \,\big] + X,$$

we deduce

$$Z = V - \mathbb{E}\big[\, V \,\|\, X \,\big] = Y - \mathbb{E}\big[\, Y \,\|\, X \,\big]$$

so Z is $\boldsymbol{Y}$-valued and its distribution is the centered Gaussian measure on $\boldsymbol{Y}$ with variance operator $\Delta_{Y,X}$. The equality (1.1.22) now follows from the regression formula (1.1.19).

To prove (1.1.23) observe that, since $\Gamma_{\Delta_{X,Y}}$ is nondegenerate, we have $\Gamma_{\Delta_{X,Y}}\big[\, \mathcal{O} \,\big] > 0$, for any open subset $\mathcal{O}$ of $\ker L$. Choose $c > 0$ such that the open set $\big\{ f \,|_{\ker L} > c \big\}$ is nonempty. Then

$$\int_{\ker L} f(y)\Gamma_{\Delta_{Y,X}}\big[\, dy \,\big] > c\Gamma_{\Delta_{X,Y}}\big[\, \{f > c\} \cap \ker L \,\big] > 0.$$

□

Remark 1.1.36. The nondegeneracy of $\Gamma_{\Delta_{Y,X}}$ is important. If $\Gamma_{\Delta_{Y,X}}$ were concentrated on a proper subspace $Z \subset \ker L$, it would still be possible that f is nontrivial yet $f \,|_Z = 0$. □

1.1.3. Complex Gaussian variables and vectors. A complex random variable

$$Z = X + \boldsymbol{i}Y : \big(\Omega, \mathcal{S}, \mathbb{P}\big) \to \mathbb{C}$$

is called *Gaussian* if the random vector $(X, Y) = (\mathbf{Re}\, Z, \mathbf{Im}\, Z)$ is Gaussian. For simplicity in the sequel we will focus exclusively on centered variables so we will assume X, Y are centered

The variance of the complex Gaussian random variable $X + \boldsymbol{i}Y$ is viewed as a real Gaussian vectors represented by the 2×2-matrix

$$A = \mathrm{Var}_{\mathbb{R}}\big[\, Z \,\big] := \begin{bmatrix} \mathrm{Var}\big[\, X \,\big] & \mathbb{E}\big[\, XY \,\big] \\ \mathbb{E}\big[\, XY \,\big] & \mathrm{Var}\big[\, Y \,\big] \end{bmatrix}.$$

However, in this book we will work with a more restrictive concept of complex Gaussian random variable.

Definition 1.1.37. The complex random variable Z is *symmetric* if the random variables Z and $\boldsymbol{i}Z$ have the same distribution. □

This means that $\operatorname{Var}_{\mathbb{R}}\big[\,Z\,\big] = \operatorname{Var}_{\mathbb{R}}\big[\,\boldsymbol{i}Z\,\big]$. It is well known that the multiplication by $\boldsymbol{i}$, viewed as a real linear operator $\mathbb{R}^2 \to \mathbb{R}^2$ is represented by the 2×2-matrix

$$J = \left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right].$$

Using (1.1.10) we deduce that Z is symmetric iff

$$A = -JAJ \Longleftrightarrow JA = AJ$$

The only symmetric 2×2 matrices that commute with J are the scalar multiples of the identity. Hence Z is symmetric iff X, Y are i.i.d. normal random variables. In this case

$$\operatorname{Var}_{\mathbb{R}}\big[\,Z\,\big] = v\mathbb{1}_{\mathbb{R}^2},\ \ v = \operatorname{Var}\big[\,X\,\big] = \operatorname{Var}\big[\,Y\,\big].$$

An elementary computation shows that Z is symmetric if and only if $\mathbb{E}\big[\,Z^2\,\big] = 0$. In this case

$$v = \frac{1}{2}\mathbb{E}\big[\,Z\bar{Z}\,\big].$$

Proposition 1.1.38. *Suppose that $Z = (Z_1, \dots, Z_n) : \big(\Omega, \mathcal{S}, \mathbb{P}\big) \to \mathbb{C}^n$ is a centered random vector satisfying the following condition:*

for any $u_1, \dots, u_n \in \mathbb{C}$ the complex random variable

$$u_1 Z + u_2 Z_2 + \cdots + u_n Z_n \tag{C}$$

a symmetric complex Gaussian random variables.

Denote by $\operatorname{Var}_{\mathbb{C}}\big[\,Z\,\big]$ *the* complex variance matrix *of Z, i.e., the $n\times n$ hermitian matrix*

$$\operatorname{Var}_{\mathbb{C}}\big[\,Z\,\big] := \Big(\mathbb{E}\big[\,Z_j\bar{Z}_k\,\big]\Big)_{1\le j,k\le n}.$$

Note that $\operatorname{Var}_{\mathbb{C}}\big[\,Z\,\big]$ *can be viewed either as a* complex linear *operator $\mathbb{C}^n \to \mathbb{C}^n$, or as a* real linear *operator $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$.*

Then the following hold

(i) *The random vector Z, viewed as a real random vector $\big(\Omega, \mathcal{S}, \mathbb{P}\big) \to \mathbb{R}^{2n} \cong \mathbb{C}^n$, is Gaussian.*

(ii) $\mathbb{E}\big[\,Z_j Z_k\,\big] = 0$, $\forall j, k = 1, \dots, n$.

(iii) *Denote by* $\operatorname{Var}_{\mathbb{R}}\big[\,Z\,\big]$ *the variance operator of Z viewed as a* real *Gaussian vectors. Then*

$$\operatorname{Var}_{\mathbb{R}}\big[\,Z\,\big] = \frac{1}{2}\operatorname{Var}_{\mathbb{C}}\big[\,Z\,\big].$$

Above, both sides are viewed as real linear *operators $\mathbb{R}^n \to \mathbb{R}^n$.*

Proof. (i) Set $X_k := \mathbf{Re}\, Z_k$ and $Y_k = \mathbf{Im}\, Z_k$. For any complex numbers $u_k = s_k - \boldsymbol{i} t_k$ the random variable

$$\sum_k u_k Z_k = \sum_k \left(s_k X_k + t_k Y_k \right) + \boldsymbol{i} \sum_k \left(s_k Y_k - t_k X_k \right)$$

is a symmetric complex Gaussian variable. This implies that the real random vector

$$\left(X_1, Y_1, \dots, X_n, Y_n \right)$$

is Gaussian. and that each of the complex Gaussian variables Z_k is symmetric.

(ii) Set $v_k = \mathbb{E}\left[X_k^2 \right] = \mathbb{E}\left[Y_k^2 \right]$. Suppose that $u_k = 0$ for $k \neq 1, 2$ $u_j = s_j - \boldsymbol{i} t_j$, $j = 1, 2$. The $u_1 Z_1 + u_2 Z_2 = A(s,t) + \boldsymbol{i} B(s,t)$ is a symmetric complex Gaussian variable. We deduce that real Gaussian random variables

$$A(s,t) = s_1 X_1 + t_1 Y_1 + s_2 X_2 + t_2 Y_2 \;\text{ and }\; B(s,t) = s_1 Y_1 - t_1 X_1 + s_2 Y_2 - t_2 X_2$$

are i.i.d.. Suppose that $t_1 = s_2 = 0$. Then

$$\begin{aligned} \operatorname{Var}\left[A \right] &= s_1^2 v_1 + t_2^2 v_2 + 2 s_1 t_2 \mathbb{E}\left[X_1 Y_2 \right] \\ = \operatorname{Var}\left[B \right] &= s_1^2 v_1 + t_2^2 v_2 - 2 s_1 t_2 \mathbb{E}\left[Y_1 X_2 \right]. \end{aligned}$$

From the equality $\operatorname{Var}\left[A \right] = \operatorname{Var}\left[B \right]$ we deduce that

$$\mathbb{E}\left[X_1 Y_2 \right] = -\mathbb{E}\left[Y_1 X_2 \right].$$

From the equality $\mathbb{E}\left[AB \right] = 0$ we deduce that

$$\mathbb{E}\left[X_1 X_2 \right] = \mathbb{E}\left[Y_1 Y_2 \right].$$

We can rewrite these equalities compactly as $\mathbb{E}\left[Z_1 Z_2 \right] = 0$. Clearly the above argument shows that $\mathbb{E}\left[Z_j Z_j \right] = 0$ for any j, k.

(iii) For $j \neq k$ We set

$$a_{jk} := \mathbb{E}\left[X_j X_k \right],\;\; b_{jk} := -\mathbb{E}\left[X_j Y_k \right],\;\; z_{jk} = a_{jk} + \boldsymbol{i} b_{jk} = \frac{1}{2} \mathbb{E}\left[Z_j \bar{Z}_k \right]$$

The covariance operator of the two variables Z_j, Z_j viewed as two-dimensional random vectors is

$$\operatorname{Cov}[Z_j, Z_k] = \begin{bmatrix} \mathbb{E}\left[X_j X_k \right] & \mathbb{E}\left[X_j Y_k \right] \\ \mathbb{E}\left[Y_j X_k \right] & \mathbb{E}\left[Y_k Y_k \right] \end{bmatrix} = \begin{bmatrix} a_{jk} & -b_{jk} \\ b_{jk} & a_{jk} \end{bmatrix}.$$

The above matrix describes the multiplication by z_{jk} viewed as a real linear operator $\mathbb{C} \to \mathbb{C}$. □

Definition 1.1.39. A complex random vector $Z : \left(\Omega, \mathcal{S}, \mathbb{P} \right) \to \mathbb{C}^n$ is called a *symmetric complex Gaussian vector* if it satisfies condition (**C**) in Proposition 1.1.38..

A collection $Z_1, \dots, Z_n$ of symmetric complex random variables is called *jointly Gaussian* if the random vector $(Z_1, \dots, Z_n)$ is a complex symmetric Gaussian vector. □

Remark 1.1.40. Suppose that $\boldsymbol{U}$ is a finite dimensional complex Euclidean space, $m = \dim_{\mathbb{C}} \boldsymbol{U}$. Denote by $\langle -, - \rangle$ the associated Hermitian[1] inner product. This defines a *real* inner product.

$$\left(\boldsymbol{u}_1, \boldsymbol{u}_2 \right) = \mathbf{Re} \left\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \right\rangle.$$

[1]We adhere to the geometers' convention that a Hermitian inner product is conjugate linear in the second variable. Physicists a ssume that the Hermitian inner product is conjugate linear in the first variable.

Suppose that Z is a $\boldsymbol{U}$-valued centered random vector that is Gaussian when we view $\boldsymbol{U}$ as a *real* vector space. As such Z has a variance $A = \operatorname{Var}_{\mathbb{R}}\big[\,Z\,\big]$ which is a symmetric real linear operator $\boldsymbol{U} \to \boldsymbol{U}$. Then Z is a symmetric complex Gaussian iff $\operatorname{Var}_{\mathbb{R}}\big[\,Z\,\big]$ is also complex linear, i.e.,

$$A(\boldsymbol{i}\boldsymbol{u}) = \boldsymbol{i}A\boldsymbol{u}, \ \ \forall \boldsymbol{u} \in \boldsymbol{U}.$$

We will denote by $\operatorname{Var}_{\mathbb{C}}\big[\,Z\,\big]$ This complex linear operator.

If $\big(\,\boldsymbol{e}_j\,\big)_{1\leq j\leq m}$ we fix a complex orthonormal basis of $\boldsymbol{V}$, then $\operatorname{Var}_{\mathbb{C}}\big[\,Z\,\big]$ is represented in this basis by a complex Hermitian $m \times m$ matrix. The collection

$$\boldsymbol{e}_1, \boldsymbol{i}\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{i}\boldsymbol{e}_m$$

is a real orthonormal basis of $\boldsymbol{U}$. In this basis, $\operatorname{Var}_{\mathbb{R}}\big[\,Z\,\big]$ is represent by a real symmetric $2m \times 2m$-matrix □

1.1.4. Gaussian measures on Fréchet spaces. In this brief, mostly expository, subsection I want to describe a few facts about Gaussian measures on infinite dimensional spaces. For more details I refer to [**21, 148**].

Let me first recall a rather deep and very versatile measurability result. I will state only a special case that suffices for all the applications I have in mind. For a proof and a more general version I refer to [**36**, Sec. 8.6].

Theorem 1.1.41 (Blackwell)**.** *Suppose that X is a Polish space, i.e., a complete separable metric space. Denote by $\mathcal{B}_X$ the sigma-algebra of Borel subsets of X. Suppose that $\mathfrak{F}$ is a countable family of Borel measurable functions on X that separates points. Then $\sigma(\mathfrak{F}) = \mathcal{B}_X$, where $\sigma(\mathfrak{F})$ denotes the sigma-algebra generated by $\mathfrak{F}$.* □

Digression 1.1.42. I want to digress to discuss an infinite dimensional version of the Cramér-Wold theorem, [**80**, Cor. 6.5].

Let Ω be a set and V a vector spaces of functions $f : \Omega \to \mathbb{R}$. Denote by $\sigma(V)$ the sigma-algebra generated by the collection V. Note that

$$\sigma(V) = \bigvee_{F \in 2_0^V} \sigma(F),$$

where we recall that 2_0^S denotes the collection of finite subsets of a set S. For any $F \in 2_0^V$ we have a natural linear map

$$p_F : F \to \mathbb{R}^F, \ \ \Omega \ni \omega \mapsto \big(\,f(\omega)\,\big)_{f\in F} \in \mathbb{R}^F.$$

Let μ be a probability measure on $\big(\,\Omega, \sigma(V)\,\big)$. For any complex valued, bounded measurable function φ we denote by $\mu\big[\,\varphi\,\big]$ the integral of φ with respect to μ.

The measure μ is uniquely determined by the marginals $\mu^F = (p_F)_{\#}\mu \in \operatorname{Prob}\big(\,\mathbb{R}^F\,\big)$. Indeed, the marginal μ^F determines the restriction of μ on $\sigma(F)$ and the collection $\big(\,\sigma(F)\,\big)_{F\in 2_0^V}$ is a π-system that generates $\sigma(V)$.

The probability measure μ^F on $\mathbb{R}^F$ is uniquely determined by its Fourier transform

$$\widehat{\mu}^F : \operatorname{span}(F) \to \mathbb{C}, \ \ g \mapsto \mu\big[\,e^{\boldsymbol{i}g}\,\big].$$

Equivalently the measure μ^F is uniquely determined by the distributions of the random variables $g : \big(\,\Omega, \sigma(V)\,\big) \to \mathbb{R}$, $g \in V$. Summarizing, we deduce that following result.

Proposition 1.1.43. *Let Ω be a set, V a vector space of real valued functions on Ω. Then a probability probability measure μ on $\sigma(V)$ is uniquely determined by either one of the following data.*

(i) *The distributions of the collection of random variables $f : \Omega \to \mathbb{R}$, $f \in V$.*

(ii) *The distributions of the collection of complex random variables $e^{\boldsymbol{i}f}$, $f \in V$.*

(iii) *The Fourier transform*

$$\widehat{\mu} : V \to \mathbb{C}, \;\; \widehat{\mu}\big[\, f \,\big] = \mu\big[\, e^{\boldsymbol{i}f} \,\big].$$

□

For a different proof of this proposition we refer to [46, Sec. 8.1].

To see this principle at work, consider two sets Ω_0, Ω_1 and two vector spaces $V_i \subset \mathbb{R}^{\Omega_i}$, $i = 0, 1$. These vector spaces determine two sigma-algebras $\sigma(V_i)$, $i = 0, 1$. Consider the product space $\Omega = \Omega_0 \times \Omega_1$ equipped with the product sigma-algebra $\sigma(V_0) \otimes \sigma(V_1)$. Let $p_i : \Omega_0 \times \Omega_1 \to \Omega_i$ denote the canonical projection. Note that

$$\sigma(V_0) \otimes \sigma(V_1) = \sigma(V_0 \boxplus V_1)$$

where $V_0 \boxplus V_1$ denotes the space $p_0^* V_0 + p_1^* V_1$ of functions $f : \Omega_0 \times \Omega_1 \to \mathbb{R}$ with the following property: $\exists f_i \in V_i$, $i = 0, 1$ so that

$$f(\omega_0, \omega_1) = f_0(\omega_0) + f_1(\omega_1), \;\; \forall (\omega_0, \omega_1) \in \Omega_0 \times \Omega_1.$$

Let ν be a probability measure on $\big(\Omega_0 \times \Omega_1, \sigma(V_0) \otimes \sigma(V_1) \big)$. Denote by ν_i the marginals of ν, $\nu_i := (p_i)_{\#}\nu$, $i = 0, 1$.

Suppose that μ_i is a probability measure on $\big(\Omega_i, \sigma(V_i) \big)$, $i = 0, 1$. To verify that $\nu = \mu_0 \times \mu_1$ it suffices to check that $\forall f_0 \in V_0$, $f_1 \in V_1$,

$$\int_{\Omega_0 \times \Omega_1} e^{\boldsymbol{i}(f_0(\omega_0) + f_1(\omega_1))} \nu\big[\, d\omega_0 d\omega_1 \,\big] = \left(\int_{\Omega_0} e^{\boldsymbol{i}f_0(\omega_0)} \mu_0\big[\, d\omega_0 \,\big] \right) \left(\int_{\Omega_1} e^{\boldsymbol{i}f_1(\omega_1)} \mu_1\big[\, d\omega_1 \,\big] \right).$$

This completes the digression. □

Recall that a *Fréchet space* is a vector space $\boldsymbol{X}$ equipped with a countable family of seminorms

$$\| - \|_\nu : X \to [0, \infty), \;\; \nu \in \mathbb{N},$$

such that the function

$$d : \boldsymbol{X} \times \boldsymbol{X} \to [0, \infty), \;\; d(x_0, x_1) := \sum_{\nu \in \mathbb{N}} \frac{1}{2^\nu} \max\big(\|x_0 - x_1\|_\nu, \, 1 \big)$$

defines a *complete metric* on $\boldsymbol{X}$. Note that the metric d is translation invariant. A subset $S \subset \boldsymbol{X}$ is said to be *bounded* if

$$\sup_{s \in S} \|s\|_\nu < \infty, \;\; \forall \nu \in \mathbb{N}.$$

Example 1.1.44. (a) Banach spaces are Fréchet spaces.

(b) Suppose that $U \subset \mathbb{R}^d$ is an open set. Then for any $m \geq 0$ the space $C^m(U)$ is a separable Fréchet space. To see this choose a compact exhaustion $(K_\nu)_{\nu \geq 1}$ i.e. countable family of of compact subsets $K|nu \subset U$ such that

$$\forall \nu \geq 1, \;\; K_\nu \subset \boldsymbol{int}(K_{\nu+1}) \;\text{ and } \bigcup_{\nu \geq 1} K_\nu = U.$$

The topology is defined by the seminorms

$$\|f\|_\nu = \sup_{x \in K_\nu} \big(|f(x)| + |Df(x)| + \cdots + |D^m(x)| \big)$$

The topology determined by these seminorms corresponds to uniform convergence on compacts.

To see that $C^m(U)$ separable note that the space of polynomials in d variables with rational coefficients is dense in $C^m(U)$, [**154**, Chap. 15, Cor. 4]. □

Let $\boldsymbol{X}$ be a real *separable* Fréchet space. In this case the Borel sigma-algebra of $\boldsymbol{X} \times \boldsymbol{X}$ coincides[2]with the product of sigma-algebras $\mathcal{B}_{\boldsymbol{X}}$

$$\mathcal{B}_{\boldsymbol{X} \times \boldsymbol{X}} = \mathcal{B}_{\boldsymbol{X}} \otimes \mathcal{B}_{\boldsymbol{X}}.$$

Since the addition $+ : X \to X \to X$ is continuous it is $\mathcal{B}_{\boldsymbol{X} \times \boldsymbol{X}}$=measurable.

We denote by $\boldsymbol{X}^*$ the topological dual of $\boldsymbol{X}$ and by

$$\langle -, - \rangle : \boldsymbol{X}^* \times \boldsymbol{X} \to \mathbb{R}$$

the natural pairing

$$\boldsymbol{X}^* \times \boldsymbol{X} \ni (\xi, x) \mapsto \langle \xi, x \rangle := \xi(x).$$

The dual $\boldsymbol{X}^*$ is equipped with several useful topologies

$$\sigma\big(\boldsymbol{X}^*, \boldsymbol{X}\big) \subset \tau\big(\boldsymbol{X}^*, \boldsymbol{X}\big) \subset \beta\big(\boldsymbol{X}^*, \boldsymbol{X}\big). \tag{1.1.24}$$

- The topology $\sigma\big(\boldsymbol{X}^*, \boldsymbol{X}\big)$, also know as the *weak* topology*, corresponds to the uniform convergence on the finite subsets of X.
- The topology $\tau\big(\boldsymbol{X}^*, \boldsymbol{X}\big)$, also known as the *Mackey topology,* corresponds to uniform convergence on the symmetric, compact convex subsets of $\boldsymbol{X}$.
- The topology $\beta\big(\boldsymbol{X}^*, \boldsymbol{X}\big)$, also known as the *strong topology*, corresponds to uniform convergence on the bounded subsets of $\boldsymbol{X}$.

For $a \in \{\sigma, \tau, \beta\}$ we denote by $\boldsymbol{X}^*_a$ the dual equipped with the $a\big(\boldsymbol{X}^*, \boldsymbol{X}\big)$-topology. The Mackey-Arens theorem shows that for $a = \sigma, \tau$, the topological dual of $\boldsymbol{X}^*_a$ can be identified with $\boldsymbol{X}$; see [**140**, Sec.IV.3]. This means that a linear function $L : \boldsymbol{X}^* \to \mathbb{R}$ is $a\big(\boldsymbol{X}^*, \boldsymbol{X}\big)$-continuous iff there exists $x \in \boldsymbol{X}$ such that $L(\xi) = \langle \xi, x \rangle$, $\forall \xi \in \boldsymbol{X}^*$.

Proposition 1.1.45. *The Borel sigma-algebra of $\boldsymbol{X}$ coincides with the sigma-algebra $\sigma(\boldsymbol{X}^*)$ generated by the family of continuous linear functions $\xi : \boldsymbol{X} \to \mathbb{R}$.*

[2]There is this the strange Nedoma pathology: if X is a metric space, then the diagonal $\Delta \subset X \times X$ is closed in the product topology and thus Borel measurable in this topology. However, if the cardinality of X is bigger than the cardinality of the continuum, then Δ *does not* belong to the sigma-algebra $\mathcal{B}_X \otimes \mathcal{B}_X$, so $\mathcal{B}_X \otimes \mathcal{B}_X \subsetneq \mathcal{B}_{X \times X}$.

Proof. Let $\{x_n\}_{n\in\mathbb{N}}$ be a countable dense subset of $\boldsymbol{X}$. We assume $x_n \neq 0$, $\forall n$. We deduce from the Hahn-Banach theorem that for any $n \in \mathbb{N}$ there exists $\xi_n \in X^*$ such that $\xi_n(x_n) = 1$. The collection $\Xi = \{\xi_n\}_{n\in\mathbb{N}} \subset C(\boldsymbol{X})$ separates the points. We have $\sigma(\Xi) \subset \sigma(\boldsymbol{X}^*) \subset \mathcal{B}_{\boldsymbol{X}}$. Blackwell's theorem now shows that $\sigma(\Xi^*) = \mathcal{B}_{\boldsymbol{X}}$. □

The Fourier transform of a Borel probability measure $\mu \in \operatorname{Prob}(\boldsymbol{X})$ is the function

$$\widehat{\mu} : \boldsymbol{X}^* \to \mathbb{C},\ \ \widehat{\mu}(\xi) = \mathbb{E}\big[\, e^{\boldsymbol{i}\xi}\,\big].$$

Since $\mathcal{B}_{\boldsymbol{X}} = \sigma(\boldsymbol{X}^*)$ we deduce from Proposition 1.1.43 that μ is uniquely determined by its Fourier transform. More generally, we have the following result.

Corollary 1.1.46. *Suppose that $\boldsymbol{X}$ is a separable Fréchet space and $\mathcal{L} \subset \boldsymbol{X}^*$ is a subspace such that $\sigma(\mathcal{L}) = \sigma(\boldsymbol{X}^*)$. Let $\mu_0, \mu_1 \in \operatorname{Prob}(\boldsymbol{X})$. Then*

$$\mu_0 = \mu_1 \Longleftrightarrow \widehat{\mu_0}(\xi) = \widehat{\mu_1}(\xi),\ \ \forall \xi \in \mathcal{L}.$$

□

Definition 1.1.47. A Borel probability measure Γ on the separable Fréchet space $\boldsymbol{X}$ is called *Gaussian* if any continuous linear functional $\xi \in \boldsymbol{X}^*$, viewed as a random variable, is Gaussian. Equivalently,

$$\xi_\#\Gamma = \gamma_{m[\xi],v[\xi]},\ \ \forall \xi \in \boldsymbol{X}^*.$$

The Gaussian measure is called *nondegenerate* if $v[\xi] > 0$, $\forall \xi \in \boldsymbol{X}^* \setminus \{0\}$. It is called *centered* if $m[\xi] = 0$, $\forall \xi \in \boldsymbol{X}^*$. □

We see that $\Gamma \in \operatorname{Prob}(X)$ is centered Gaussian if, $\forall \xi \in \boldsymbol{X}^*$

$$\widehat{\Gamma}(\xi) = e^{-v[\xi]/2},\ \ v[\xi] = \mathbb{E}_\Gamma\big[\,\xi\,\big] = \int_X \xi(x)^2\Gamma\big[\,dx\,\big]$$

We deduce that a centered Gaussian measure is uniquely determined by the variance function

$$\operatorname{Var} : \boldsymbol{X}^* \to \mathbb{R},\ \ \xi \mapsto \mathbb{E}_\Gamma\big[\,\xi^2\,\big].$$

Note that $\forall t \in \mathbb{R},\ \ \forall \xi, \eta \in X^*$

$$\operatorname{Var}[t\xi] = t^2 \operatorname{Var}[\xi],\ \ \operatorname{Var}[\xi+\eta] + \operatorname{Var}[\xi-\eta] = 2\big(\operatorname{Var}[\xi] + \operatorname{Var}[\eta]\,\big). \tag{1.1.25}$$

Proposition 1.1.48. *Let $\boldsymbol{X}$ be a separable Fréchet space and $\mu \in \operatorname{Prob}(\boldsymbol{X})$. Denote by R the "rotation"*

$$R : X \times \boldsymbol{X} \to X \times \boldsymbol{X},\ \ R(x_0, x_1) = \Big(\frac{1}{\sqrt{2}}(x_0 - x_1), \frac{1}{\sqrt{2}}(x_0 + x_1)\Big).$$

Then the following are equivalent.

(i) *The measure μ is Gaussian.*

(ii) $\mu \otimes \mu = R_\#(\mu \times \mu)$

Proof. For $i = 0, 1$ we denote by p_i the natural projection $\boldsymbol{X} \times \boldsymbol{X} \ni (x_0, x_1) \mapsto x_i \in \boldsymbol{X}$. If $\xi \in \boldsymbol{X}^*$ and $i = 0, 1$, then we set $\xi_i = \xi \circ p_i \in (\boldsymbol{X} \times \boldsymbol{X})^*$. Set $\nu := R_\#(\mu \otimes \mu)$. Note that

$$\xi_0 \circ R = \frac{1}{\sqrt{2}}\big(\,\xi_0 - \xi_1\,\big),\ \ \xi_1 \circ R = \frac{1}{\sqrt{2}}\big(\,\xi_0 + \xi_1\,\big)$$

(i) $\Rightarrow$ (ii) To show $\mu\otimes\mu=\nu$ we use Corollary 1.1.46. We have to show that if μ is Gaussian then

$$\forall\xi,\eta\in\boldsymbol{X}^*:\quad \int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\boldsymbol{i}(\xi_0+\eta_1)}d\nu=\left(\int_{\boldsymbol{X}} e^{\boldsymbol{i}\xi}d\mu\right)\left(\int_{\boldsymbol{X}} e^{\boldsymbol{i}\eta}d\mu\right). \tag{1.1.26}$$

Proof of (1.1.26) We have

$$\begin{aligned}\int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\boldsymbol{i}(\xi_0+\eta_1)}d\nu&=\int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\boldsymbol{i}(\xi_0\circ R(x_0,x_1)+\eta_1\circ R(x_0,x_1))}\mu\otimes\mu\big[\,dx_0dx_1\,\big]\\&=\int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\frac{\boldsymbol{i}}{\sqrt{2}}((\xi+\eta)(x_0)-(\xi-\eta)(x_1))}\mu\otimes\mu\big[\,dx_0dx_1\,\big]\\&=\left(\int_{\boldsymbol{X}} e^{\frac{\boldsymbol{i}}{\sqrt{2}}(\xi+\eta)}d\mu\right)n\left(\int_{\boldsymbol{X}} e^{\frac{-\boldsymbol{i}}{\sqrt{2}}(\xi-\eta)}d\mu\right)\\&=e^{-v[\xi+\eta]/4}e^{-v[\xi-\eta]/4}\stackrel{(1.1.25)}{=}e^{-v[\xi]/2-v[\eta/2]}=\left(\int_{\boldsymbol{X}} e^{\boldsymbol{i}\xi}d\mu\right)\left(\int_{\boldsymbol{X}} e^{\boldsymbol{i}\eta}d\mu\right).\end{aligned}$$

(ii) $\Rightarrow$ (i) To show that μ is Gaussian it suffices to show that for any $\xi\in\boldsymbol{X}^*$, the random variable $\xi:(\boldsymbol{X},\mu)\to\mathbb{R}$ is Gaussian. Note that the random variables

$$\xi_0,\xi_1:(\boldsymbol{X}\times\boldsymbol{X},\mu\otimes\mu)\to\mathbb{R}$$

are independent copies of $\boldsymbol{X}$, i.e., they are independent and they have the same distribution as ξ. According to Polya's Theorem 1.1.8 it suffices to show that the random variables ξ and $\alpha=\frac{1}{\sqrt{2}}(\xi_1+\xi_0)$ have the same distribution, i.e.,

$$\mathbb{E}\big[\,e^{\boldsymbol{i}t\alpha}\,\big]=\mathbb{E}\big[\,e^{\boldsymbol{i}t\xi}\,\big],\ \ \forall t\in\mathbb{R}.$$

We have

$$\mathbb{E}\big[\,e^{\boldsymbol{i}t\alpha}\,\big]=\int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\frac{\boldsymbol{i}t}{\sqrt{2}}(\xi_0+\xi_1)(x_0,x_1)}\mu\otimes\mu\big[\,dx_0dx_1\,\big]=\int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\boldsymbol{i}t\xi_1\circ R(x_0,x_1)}\mu\otimes\mu\big[\,dx_0dx_1\,\big]$$

$(\nu=R_\#(\mu\otimes\mu))$

$$=\int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\boldsymbol{i}t\xi(x_1)}\nu\big[\,dx_0dx_1\,\big]$$

$(\nu=\mu\otimes\mu)$

$$=\int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\boldsymbol{i}t\xi(x_1)}\mu\otimes\mu\big[\,dx_0dx_1\,\big]=\int_X e^{\boldsymbol{i}t\xi(x)}\mu\big[\,dx\,\big]=\mathbb{E}\big[\,e^{\boldsymbol{i}t\xi}\,\big].$$

□

Corollary 1.1.49. *Suppose that $\boldsymbol{X}$ is a separable Fréchet space, $\mu\in\mathrm{Prob}(\boldsymbol{X})$ and $\mathcal{L}\subset\boldsymbol{X}^*$ is a subspace such that $\sigma(\mathcal{L})=\mathcal{B}_{\boldsymbol{X}}$. Then the following are equivalent.*

(i) *The measure μ is centered Gaussian,*

(ii) *For any $\xi\in V$ the random variable $\xi:(\boldsymbol{X},\mu)\to\mathbb{R}$ is centered Gaussian.*

Proof. Clearly (i) $\Rightarrow$ (ii) so it suffices to prove (ii) $\Rightarrow$ (i). For each $\xi\in\mathcal{L}$ we denote by $v[\xi]$ the variance of ξ. It satisfies the equalities (1.1.48). Using Proposition 1.1.48 it suffices to show that $\mu\otimes\mu=\nu=R_\#(\mu\otimes\mu)$. Since $\mathcal{B}_{\boldsymbol{X}\times\boldsymbol{X}}=\mathcal{B}_{\boldsymbol{X}}\otimes\mathcal{B}_{\boldsymbol{X}}=\sigma(\mathcal{L})\otimes\sigma(\mathcal{L})$ we can use the strategy outlined in Digression 1.1.42 so it suffices to show that

$$\forall\xi,\eta\in\mathcal{L}:\quad \int_{\boldsymbol{X}\times\boldsymbol{X}} e^{\boldsymbol{i}(\xi_0+\eta_1)}d\nu=\left(\int_{\boldsymbol{X}} e^{\boldsymbol{i}\xi}d\mu\right)\left(\int_{\boldsymbol{X}} e^{\boldsymbol{i}\eta}d\mu\right). \tag{1.1.27}$$

Since $\xi+\eta, \xi-\eta \in \mathcal{L}$, we deduce that $\xi+\eta, \xi-\eta$ are Gaussian variables as well and the proof of (1.1.26) carries over with no modification to this situation as well. □

Remark 1.1.50. The Corollaries 1.1.46 and 1.1.49 may suggest that a Gaussian measure Γ is nondegenerate iff ξ is a nondegenerate Gaussian random variable for any $\xi \in \mathcal{L}$. Example 1.2.20 will show that this is not the case. □

The *covariance form* of a centered Gaussian measure Γ on a separable Freéchet space $\boldsymbol{X}$ is the continuous, symmetric bilinear form

$$C_\Gamma : X^* \times X^* \to \mathbb{R},\ \ C_\Gamma(\xi,\eta) = \mathbb{E}_\Gamma\big[\, \xi \cdot \eta \,\big] = \int_X \xi(x)\eta(x)\Gamma\big[\, dx \,\big].$$

Note that each $\xi \in \boldsymbol{X}^*$ is a function on $\boldsymbol{X}$ that is L^2 with respect to the measure Γ. This determines a tautological linear map

$$T_\Gamma : \boldsymbol{X}^* \to L^2(\boldsymbol{X},\Gamma) \tag{1.1.28}$$

that associates to each continuous linear functional $\xi : \boldsymbol{X} \to \mathbb{R}$ its Γ-a.s. equivalence class. The map T_Γ induces a continuous map $X^*_\tau \to L^2(\boldsymbol{X},\Gamma)$; see [**21**, Lemma 3.2.1] or [**155**, Thm.3(3)]. As such, it has a continuous dual map

$$T^*_\Gamma : L^2(\boldsymbol{X},\Gamma) \to (\boldsymbol{X}^*_\tau)^* = \boldsymbol{X}.$$

More precisely $T^*_\Gamma \xi$ is the linear functional u on $\boldsymbol{X}^*$ such that

$$u(\eta) = \mathbb{E}\big[\, \xi\eta \,\big],\ \ \forall \eta \in \boldsymbol{X}^*. \tag{1.1.29}$$

We denote by R_Γ the composition

$$R_\Gamma := T^*_\Gamma T_\Gamma : \boldsymbol{X}^* \to \boldsymbol{X}. \tag{1.1.30}$$

The map $R_\Gamma : \boldsymbol{X}^* \to \boldsymbol{X}$ is uniquely determined by the conditions

$$\langle \eta, R_\Gamma \xi \rangle = C_\Gamma(\xi,\eta) = \mathbb{E}_\Gamma\big[\, \xi\eta \,\big],\ \ \forall \xi,\eta \in \boldsymbol{X}^*.$$

Note that $\ker R_\Gamma = \ker T_\Gamma$ and these maps are injective iff Γ is nondegenerate.

For a proof of the following fundamental fact we refer to [**47**, Sec. 3], [**60**, Sec.1] or [**148**, Sec. 3.2.2].

Theorem 1.1.51 (Fernique)**.** *Let Γ be a centered Gaussian measure on the separable Fréchet space $\boldsymbol{X}$ defined by a sequence of seminorms $(\|-\|_\nu)_{\nu\geq 0}$. Fix $\nu \geq 0$ $r_0 = r_0(\nu) > 0$ such that*

$$\Gamma\big[\, \{\|x\|_\nu \leq r_0\} \,\big] = q > \frac{1}{2}.$$

Set

$$A := A(r_0,q) = \frac{1}{24r_0^2}\log\Big(\, \frac{q}{1-q} \,\Big).$$

Then, for any $r > r_0$ we have Fernique's inequality

$$\Gamma\big[\, \{\|x\|_\nu > r\} \,\big] \leq r_0 e^{-Ar^2}. \tag{1.1.31}$$

In particular

$$\int_{\boldsymbol{X}} e^{\alpha\|x\|_\nu^2}\Gamma\big[\, dx \,\big] < \infty,\ \ \forall \alpha < A, \tag{1.1.32}$$

and

$$\int_{\boldsymbol{X}} \|x\|_\nu^2 \Gamma\big[\, dx\,\big] < \infty. \tag{1.1.33}$$

□

Condition (1.1.33) implies that the map

$$T_\Gamma : \boldsymbol{X}^* \to L^2(\boldsymbol{X}, \Gamma)$$

is continuous with respect to the weak* topology on $\boldsymbol{X}^*$. The dual $T_\Gamma^* : L^2(\boldsymbol{X}, \Gamma) \to \boldsymbol{X}$ is continuous with respect to the weak topology on X. The closed graph theorem [**154**, Chap. 17, Cor.6] implies that it is also continuous with respect to original strong topology on X. We set

$$H_\Gamma^0 := R_\Gamma(\boldsymbol{X}^*) = T_\Gamma^* T_\Gamma(\boldsymbol{X}^*) \subset \boldsymbol{X}.$$

The space H_Γ^0 is a pre-Hilbert space with respect to the inner product

$$\big(\, R_\Gamma\xi, R_\Gamma\eta\,\big)_\Gamma = \mathbb{E}_\Gamma\big[\,\xi\eta\,\big],\ \ \forall \xi,\eta \in \boldsymbol{X}^*.$$

The operator T_Γ^* defines an isometry

$$T_\Gamma^* : T_\Gamma(\boldsymbol{X}^*) \subset L^2(X,\Gamma) \to H_\Gamma^0$$

and thus extends by continuity to $\boldsymbol{X}_\Gamma^*$, the closure in $L^2(\boldsymbol{X}, \Gamma)$ of $T_\Gamma(\boldsymbol{X}^*)$. We denote by H_Γ the image of this extension

$$H_\Gamma = T_\Gamma^*\big(\, \boldsymbol{X}_\Gamma^*\,\big) \subset \boldsymbol{X}. \tag{1.1.34}$$

The resulting map $T_\Gamma^* : \boldsymbol{X}_\Gamma^* \to H_\Gamma$ is a surjective isometry so H_Γ is the completion of H_Γ^0 with respect to the norm $\|-\|_\Gamma$ induced by the inner product $(-,-)_\Gamma$. The Hilbert space H_Γ is called the *Cameron-Martin space* associated to the Gaussian measure Γ.

We have the following result, [**21**, Prop. 3.1.9].

Proposition 1.1.52. *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence in $\boldsymbol{X}^*$ that separates the points in X. Then they span a dense subspace of $\boldsymbol{X}_\Gamma^*$, so $\boldsymbol{X}_\Gamma^*$ is separable. In particular, the span of the family $\big(\, R_\Gamma(\xi_n)\,\big)_{n\in\mathbb{N}}$ is dense in H_Γ.* □

For a proof of the following nontrivial result we refer to [**21**, Thm. 3.6.1].

Theorem 1.1.53 (Support theorem)**.** *Let Γ be a centered Gaussian measure on the separable Fréchet space $\boldsymbol{X}$. Then the support of Γ is the closure of H_Γ in $\boldsymbol{X}$. This means that $\Gamma\big[\, \boldsymbol{cl}(H_\Gamma)\,\big] = 1$ and for any open set U that intersects H_Γ, $\Gamma\big[\, U\,\big] > 0$.* □

Corollary 1.1.54. *Let Γ be a centered Gaussian measure on the separable Fréchet space X. Then the following are equivalent.*

(i) *The measure Γ is nondegenerate.*

(ii) *The Cameron-Martin space H_Γ is dense in $\boldsymbol{X}$.*

(iii) *For any nonempty open subset $\mathcal{O} \subset \boldsymbol{X}$, $\Gamma\big[\, \mathcal{O}\,\big] > 0$.*

□

If Γ is a finite dimensional real vector space and S is a subspace of V, then it is not hard to see that either $\Gamma\big[\,S\,\big] = 0$ or $\Gamma\big[\,S\,\big] = 1$. Xavier Fernique [60, Sec.1] proved that a similar result holds in infinite dimensions.

Theorem 1.1.55 (Zero-one law)**.** *Suppose that Γ is a Gaussian measure on the separable Fréchet space $\boldsymbol{X}$. If $\boldsymbol{Y} \subset \boldsymbol{X}$ is a Borel measurable subspace then either $\Gamma\big[\,\boldsymbol{Y}\,\big] = 0$ or $\Gamma\big[\,\boldsymbol{Y}\,\big] = 1$.* □

Proposition 1.1.56. *Suppose that $\boldsymbol{Y}, \boldsymbol{X}$ are separable Fréchet spaces and $i : \boldsymbol{Y} \to \boldsymbol{X}$ is a continuous linear injection with closed range. We have a pushforward map*

$$i_{\#} : \mathrm{Prob}(\boldsymbol{Y}) \to \mathrm{Prob}(\boldsymbol{X}).$$

(i) *A Borel probability measure $\mu \in \mathrm{Prob}(\boldsymbol{Y})$ is (centered) Gaussian if and only if its pushforward $i_{\#}\mu$ is a (centered) Gaussian probability measure on $\boldsymbol{X}$.*

(ii) *If Γ is a Gaussian probability measure on $\boldsymbol{X}$ such that $\Gamma\big[\,i(\boldsymbol{Y})\,\big] = 1$, then there exists a Gaussian measure on $\boldsymbol{Y}$ such that $\Gamma = i_{\#}\mu$.*

Proof. We have a dual map $i^* : \boldsymbol{X}^* \to \boldsymbol{Y}^*$, $\xi \mapsto \xi \circ i$. The Hahn-Banach theorem shows that this map is onto.

(i) Note that $i_{\#}\mu$ is Gaussian iff $\forall \xi \in \boldsymbol{X}^*$ the pushforward $\xi_{\#}(i_{\#}\mu) = (\xi \circ i)_{\#}\mu$ is Gaussian. Since i^* is onto, this happens iff $\eta_{\#}\mu$ is Gaussian, $\forall \eta \in \boldsymbol{Y}^*$, i. e., μ is Gaussian.

(ii) For a Borel subset $B \subset \boldsymbol{Y}$ we set

$$\mu\big[\,B\,\big] := \Gamma\big[\,i(B)\,\big].$$

Then $\mu \in \mathrm{Prob}(\boldsymbol{Y})$ and $i_{\#}\mu = \Gamma$. We deduce from (i) that μ is Gaussian. □

Theorem 1.1.53 has an immediate but useful consequence.

Proposition 1.1.57. *Let $\boldsymbol{X}$ be a separable Fréchet space. Fix a family of seminorms $(\|-\|_{\nu})_{\nu \geq 0}$ defining the topology of $\boldsymbol{X}$. Let $(x_n)_{n\geq 0}$ be a sequence in $\boldsymbol{X}$ and $(c_n)_{n\geq 0}$ a sequence of positive real numbers such that*

$$\sum_{n\geq 1} c_n \|x_n\|_{\nu} < \infty, \;\; \forall \nu.$$

Denote by $\boldsymbol{Y}$ the closure of the span of $(x_n)_{n\geq 1}$. Let $(A_n)_{n\geq 1}$ be a sequence of independent standard normal random variables defined on the probability space $(\Omega, \mathcal{S}, \mathbb{P})$. Then the following hold.

(i) *There exists a negligible subset $\mathcal{N} \in \mathcal{S}$ such that the series*

$$\sum_{n\geq 1} A_n(\omega) c_n x_n$$

converges in $\boldsymbol{X}$ to an element in $\boldsymbol{Y}$ for any $\omega \in \Omega \setminus \mathcal{N}$.

(ii) *The map $S : \Omega \to \boldsymbol{Y}$ defined by*

$$S(\omega) = \begin{cases} \sum_{n\geq 1} A_n(\omega) c_n x_n, & \omega \in \Omega \setminus \mathcal{N}, \\ 0, & \omega \in \mathcal{N} \end{cases}$$

is Borel measurable and the push-forward $\Gamma_S := S_\#\mathbb{P}$ *is a nondegenerate Gaussian measure on* $\boldsymbol{Y}$*.*

(iii) *For any nonempty open subset* $\mathcal{O} \subset \boldsymbol{Y}$, $\mathbb{P}\big[\, S \in \mathcal{O} \,\big] > 0$.

Proof. (i) We will show that the random scalar series

$$\sum_n |A_n| c_n \|x_n\|_\nu$$

is a.s. convergent for any ν. According to Kolmogorov's two-series theorem this happens if the positive random variables $X_n^\nu = |A_n| \cdot c_n \|x_n\|_\nu$ satisfy

$$\sum_{n\geq 1} \mathbb{E}\big[\, X_n^\nu \,\big] < \infty \;\text{ and }\; \sum_{n\geq 1} \mathbb{E}\big[\, (X_n^\nu)^2 \,\big] < \infty.$$

Now observe that

$$\mathbb{E}\big[\, |A_n| \,\big] = 2\frac{1}{\sqrt{2\pi}} \int_0^\infty x e^{x^2/2} dx = \sqrt{\frac{2}{\pi}},$$

$$\sum_{n\geq} \mathbb{E}\big[\, X_n^\nu \,\big] = \sqrt{\frac{2}{\pi}} \sum_{n\geq 1} c_n \|x_n\|_\nu < \infty$$

and

$$\sum_{n\geq} \mathbb{E}\big[\, (X_n^\nu)^2 \,\big] = \sum_{n\geq 1} c_n^2 \|x_n\|_\nu^2 < \infty.$$

(ii) Define $S_n : \Omega \to \boldsymbol{Y}$

$$S_n(\omega) = \begin{cases} \sum_{k=1}^n A_k(\omega) c_k x_k, & \omega \in \Omega \setminus \mathcal{N}, \\ 0, & \omega \in \mathcal{N}. \end{cases}$$

The maps S_n are measurable since the addition operation on a separable Fréchet space is a measurable map. The map S is measurable since for any $\xi \in \boldsymbol{Y}^*$ the function $\langle \xi, S \rangle$ is measurable as limit of the measurable functions $\langle \xi, S_n \rangle$.

To see that Γ_S is a Gaussian measure let $\xi \in \boldsymbol{Y}^*$. Then

$$\langle \xi, S(\omega) \rangle = \lim_{n\to\infty} \langle \xi, S_n \rangle.$$

The random variables

$$\langle \xi, S_n \rangle = \sum_{k=1}^n A_n c_n \langle \xi, x_n \rangle$$

are Gaussian as sum of independent Gaussians. Since the limit of Gaussian random variables is also Gaussian we deduce that $\langle \xi, S \rangle$ is Gaussian with variance

$$v[\xi] = \sum_{n\geq 1} c_n^2 \big|\, \langle \xi, x_n \rangle \,\big|^2.$$

Since (x_n) spans a dense subspace of $\boldsymbol{Y}$, we deduce that for any $\xi \in \boldsymbol{Y}^* \setminus 0$ such there exists n such that $\langle \xi, x_n \rangle \neq 0$. This proves that Γ_S is nondegenerate. Part (iii) now follows from Theorem 1.1.53. □

1.1.5. Mercer kernels. Classically, [**100**], a Mercer kernel on a compact interval I of the real axis is a continuous symmetric function $K : I \times I \to \mathbb{R}$ such that the associated integral operator

$$f \mapsto K[f], \;\; K[f](s) = \int_I K(s,t)f(t)dt$$

is symmetric and nonnegative definite. It this subsection we will survey some properties of a generalization of this classical concept.

Definition 1.1.58. Let $\boldsymbol{T}$ be a metric space. A *Mercer kernel* on $\boldsymbol{T}$ is a continuous function $K : \boldsymbol{T} \times \boldsymbol{T} \to \mathbb{R}$ satisfying the following properties.

(i) $K(s,t) = K(t,s)$, $\forall s,t \in \boldsymbol{T}$.

(ii) For any $t_1, \dots, t_n \in X$ the symmetric matrix $\big(K(t_i,t_j) \big)_{1 \leq i,j \leq n}$ is nonnegative definite.

□

Example 1.1.59. Let $\boldsymbol{T}$ be a compact metric space. Denote by $\boldsymbol{F}$ the Banach space $C^0(T)$ equipped with the sup norm. Suppose as in [**15**] that $\boldsymbol{U} \subset \boldsymbol{F} := C^0(\boldsymbol{T})$ is vector subspace equipped with a norm $\| - \|_U$ making it into a separable Banach space and such that the natural inclusion $\boldsymbol{U} \to \boldsymbol{F}$ is continuous. If $(t_n)_{n \in \mathbb{N}}$ is a dense subset of $\boldsymbol{T}$ the evaluation maps $\mathbf{Ev}_{t_n} \in \boldsymbol{F}^*$ separate the points in $\boldsymbol{F}$ and, according to Blackwell's Theorem 1.1.41, they generate the Borel sigma-algebra of $\boldsymbol{F}$. These evaluation maps also define continuous linear functionals on $\boldsymbol{U}$ that, a fortiori, separate the points in $\boldsymbol{U}$ so they also generate the Borel sigma algebra of $\boldsymbol{U}$.

Suppose that Γ is a centered Gaussian measure on $\boldsymbol{U}$. We deduce from Proposition 1.1.52 that the collection $\big(\mathbf{Ev}_t \big)_{t \in \boldsymbol{T}}$ spans a dense subspace of $\boldsymbol{U}^*_\Gamma$. For every $t \in \boldsymbol{T}$ we obtain a continuous function $K_t = K^\Gamma_t = R_\Gamma \mathbf{Ev}_t \in \boldsymbol{U} \subset C(\boldsymbol{T})$. The continuous function K^Γ_t is uniquely defined by the equality

$$K^\Gamma_t(s) = \mathbf{Ev}_s \big(K^\Gamma_t \big) = \int_{\boldsymbol{U}} \mathbf{Ev}_t(u) \cdot \mathbf{Ev}_s(u) \Gamma\big[du \big].$$

We set $K^\Gamma(t,s) := K^\Gamma_t(s)$. The resulting function

$$K^\Gamma : \boldsymbol{T} \times \boldsymbol{T} \to \mathbb{R}, \;\; (t,s) \mapsto K^\Gamma(x,y)$$

is called the *covariance kernel* of the Gaussian measure.

The covariance kernel $K^\Gamma : \boldsymbol{T} \times \boldsymbol{T} \to \mathbb{R}$ is a Mercer kernel. Property (ii) follows from the fact the symmetric matrix $\big(K(t_i,t_j) \big)_{1 \leq i,j \leq n}$ is the variance operator of the Gaussian vector

$$\boldsymbol{U} \to \mathbb{R}^n, \;\; u \mapsto \big(u(t_1), \dots, u(t_n) \big) \in \mathbb{R}^n.$$

The Cameron-Martin space H_Γ can be identified with the Reproducing Kernel Hilbert Space (RHKS) determined by the covariance kernel K^Γ. We refer to Appendix B.5 and the references therein for more information about this concept. □

Let us conclude with a simple way of recognizing Mercer kernels. Observe that if $K : \boldsymbol{T} \times \boldsymbol{T} \to \mathbb{R}$ is a continuous symmetric function,

$$K(s,t) = K(t,s), \;\; \forall x,y \in M,$$

then for any finite Borel measure μ on M it induces a bounded symmetric operator

$$[K] = [K]_\mu : L^2(\boldsymbol{T}, \mu) \to L^2(\boldsymbol{T}, \mu),$$
$$[K](f)(t) = \int_M K(t,s) f(s) \mu\big[\, ds\,\big]. \tag{1.1.35}$$

Note that the function K is a Mercer kernel if and only if for any $t_1, \dots, t_n \in \boldsymbol{T}$ and $\mu = \sum_i \delta_{t_i}$ the operator $[K]_\mu$ is nonnegative definite, i.e.,

$$\big(\, [K]_\mu f, f\,\big)_{L^2(\boldsymbol{T},\mu)} \geq 0, \;\; \forall f \in C^0(\boldsymbol{T}).$$

Denote by $\mathrm{Meas}(\boldsymbol{T})$ the collection of finite Borel measures on $\boldsymbol{T}$ and by $\mathrm{Prob}(\boldsymbol{T})$ the collection of Borel probability measures on M. A measure $\mu \in \mathrm{Meas}(\boldsymbol{T})$ is called *diffuse* if $\mu\big[\, U\,\big] > 0$ for any nonempty open subset $U \subset \boldsymbol{T}$

Proposition 1.1.60. *Let $K : \boldsymbol{T}\times\boldsymbol{T} \to \mathbb{R}$ be a symmetric continuous function. The following are equivalent.*

(i) *The function K is a Mercer kernel.*

(ii) *The operator $[K]_\mu$ is nonnegative definite for any $\mu \in \mathrm{Meas}(\boldsymbol{T})$.*

(iii) *The operator $[K]_\mu$ is nonnegative definite for any $\mu \in \mathrm{Prob}(\boldsymbol{T})$.*

(iv) *The operator $[K]_\mu$ is nonnegative definite for any difffuse measure probability $\mu \in \boldsymbol{T}$*

Proof. Clearly (ii) $\Rightarrow$ (i), (iii) , (iv).

Denote by $\mathcal{P}_K$ collection of measures $\mu \in \mathrm{Meas}(\boldsymbol{T})$ such that $[K]_\mu$ is nonnegative definite. Observe that if $f : M \to [0, \infty)$ is a nonnegative continuous function and $\mu \in \mathcal{P}_K$, then $f\mu \in \mathcal{P}_K$. This shows (ii) $\iff$ (iii). Hence it it suffices to show that (i) $\Rightarrow$ (iii) and (iv) $\Rightarrow$ (i).

The dominated convergence theorem shows that for any $\mu \in \mathrm{Meas}(M)$ and $f \in L^2(M)$ we the function $[K]_\mu$ is continuous and

$$\sup_{x\in M} \big|\, [K]_\mu f(x)\,\big| \leq \int_M \sup_{x,y\in M} \big|\, K(x,y)||f(y)|\mu\big[\, dy\,\big] = \|K\|_{C(M\times M}\mu\big[\, M\,\big]^{1/2}\|f\|_{L^2(M,\mu)}$$

so $[K]_\mu$ defines a continuous linear operator $L^2(\boldsymbol{T}, \mu) \to C(\boldsymbol{T})$, $\forall \mu \in \mathrm{Meas}$. We deduce from this that if $(\mu_n) \in \mathcal{P}_K$ is a sequence of measures converging weakly to a measure $\mu \in \mathrm{Meas}(M)$ then $\mu \in \mathcal{P}_K$. In other words, $\mathcal{P}_K$ is closed under the topology of weak convergence of finite measures.

Denote by $\mathrm{Prob}_0(\boldsymbol{T})$ the collection of discrete probability measures on M. More precisely, $\mu \in \mathrm{Prob}_0(\boldsymbol{T})$ iff there exist a finite set $F \subset M$ and a function $w : F \to [0, \infty)$ such that

$$\sum_{t\in F} w(t) = 1, \;\; \mu = \sum_{t\in F} w(t)\delta_t.$$

The Krein-Milman theorem shows that any μ in $\mathrm{Prob}(\boldsymbol{T})$ is the weak limit of a sequence of discrete probability measures; see [**52**, Sec. 10.1] or [**146**, Ex.8.16].

(i) $\Rightarrow$ (iii) Since K is a Mercer kernel we deduce that $\mathrm{Prob}_0(M) \in \mathcal{P}_K$ and the above discussion shows that $\mathrm{Prob}(M) \subset \mathcal{P}_K$.

(iv) ⇒ (i). Fix a non-atomic probability measure μ. We will show that $\operatorname{Prob}_0(\boldsymbol{T}) \subset \mathcal{P}_K$. Let

$$\mu_w = \sum_{t\in F} w(t)\delta_t \in \operatorname{Prob}_0(\boldsymbol{T})$$

For $t \in F$ we denote by $B_r(t)$ the open ball of radius r centered at t. For $\nu \in \mathbb{N}$, $\nu > 1/r$, choose a nonnegative continuous function $f_\nu : \boldsymbol{T} \to [0,\infty)$ with the following properties

$$\operatorname{supp} f_\nu \subset \bigcup_{t\in F} B_{1/\nu}(t), \quad \int_{B_r(t)} f_\nu(t)\mu\big[\,ds\,\big] = w(x), \quad \forall x \in F.$$

Then $f_\nu\mu \in \mathcal{P}_K$ and $f_\nu\mu$ converges weakly to μ_w as $\nu \to \infty$ so that $\mu_w \in \mathcal{P}_K$.

□

1.2. Gaussian fields

1.2.1. Random fields a.k.a. stochastic processes. This subsection has a rather modest goal namely to introduce some basic terminology and facts concerning stochastic processes. For more details we refer to two classic sources, [**45, 68**].

To put it simply, a stochastic process is a family of random quantities valued in the same measurable space. In this book I will typically use the term random maps when referring to stochastic proceses.

Definition 1.2.1. Fix a finite dimensional vector space $\boldsymbol{U}$ and a set $\boldsymbol{T}$. An $\boldsymbol{U}$-valued *random field* or *random map* on $\boldsymbol{T}$ (or parametrized by $\boldsymbol{T}$) is a map

$$X : \Omega \times \boldsymbol{T} \to \boldsymbol{U}, \quad (\omega, t) \mapsto X_\omega(t) \in \boldsymbol{U},$$

where $(\Omega, \mathcal{S}, \mathbb{P})$ is a probability space, and for any $t \in \boldsymbol{T}$, the map

$$\Omega \ni \omega \to X_\omega(t) \in \boldsymbol{U}$$

is measurable. When $\boldsymbol{U} = \mathbb{R}$, the random field X is also known as a *random function.* □

Here is an alternate viewpoint. Denote $\boldsymbol{U}^{\boldsymbol{T}}$ the space of functions $f : \boldsymbol{T} \to \boldsymbol{U}$. If

$$X : \Omega \times \boldsymbol{T} \to \boldsymbol{U}, \quad (\omega, t) \mapsto X_\omega(t) \in \boldsymbol{U},$$

then for any $\omega \in \Omega$ we have a map $X_\omega \in \boldsymbol{U}^{\boldsymbol{T}}$, $t \mapsto X_\omega(t)$. The maps X_ω are called the *sample maps* of of the random map X. We thus obtain a map

$$\Phi_X : \Omega \to \boldsymbol{U}^{\boldsymbol{T}}, \quad \Phi_X(\omega) = X_\omega.$$

For every $t \in \boldsymbol{T}$ we have a natural projection

$$\mathbf{Ev}_t : \boldsymbol{U}^{\boldsymbol{T}} \to \boldsymbol{U}, \quad f \mapsto f(t).$$

These maps determine a sigma-algebra on $\boldsymbol{U}^{\boldsymbol{T}}$, namely the smallest sigma-algebra such that all the maps $\mathbf{Ev}_t$ are Borel measurable. We denote it by $\mathcal{B}_{\boldsymbol{U}}^{\boldsymbol{T}}$. A map

$$\Psi : (\Omega, \mathcal{S}, \mathbb{P}) \to \boldsymbol{U}^{\boldsymbol{T}}, \quad \omega \mapsto \Psi_\omega$$

is measurable if and only if, for any $t \in \boldsymbol{T}$, the induced map

$$\Omega \ni \omega \mapsto \mathbf{Ev}_t(\Psi_\omega) = \Psi_\omega(t) \in \boldsymbol{U}$$

is measurable. This shows that the map Φ_X is measurable.

Conversely, any measurable map $\Phi : (\Omega, \mathcal{S}, \mathbb{P}) \to \boldsymbol{U}^{\boldsymbol{T}}$ defines a map

$$X_\Phi : \Omega \times \boldsymbol{T} \to \boldsymbol{U}, \ \ (X_\Phi)_\omega(t) - \Phi_\omega(t)$$

which is a random field in the sense of Definition 1.2.1(i). The pushforward probability measure $F_\# \mathbb{P}$ on $\boldsymbol{U}^{\boldsymbol{T}}$ is called the *distribution* of the random field X_Φ.

Denote by $2_0^{\boldsymbol{T}}$ the collection of *finite* subsets of $\boldsymbol{T}$. For any $F \in 2_0^{\boldsymbol{T}}$ we denote by π_F the natural projection $\boldsymbol{U}^{\boldsymbol{T}} \to \boldsymbol{U}^F$. Equivalently, $\pi_F(X)$ is the restriction to F of a function $X : \boldsymbol{T} \to \mathbb{R}$. Moreover, if $F_1 \subset F_2$ are two finite subsets of $\boldsymbol{T}$ we denote by $\mathcal{P}_{F_1,F_2}$ the natural projection $\boldsymbol{U}^{F_2} \to \boldsymbol{U}^{F_1}$ that maps a function $F_2 \to \boldsymbol{U}$ to its restriction to F_1.

Any probability measure μ on $\mathcal{B}_{\boldsymbol{U}}^{\boldsymbol{T}}$ determines a family of probability measures μ_F on $\mathcal{B}_{\boldsymbol{U}}^F$, $F \in 2_0^{\boldsymbol{T}}$, $\mu_F := (\pi_F)_\# \mu$. This is a *projective family* i.e., it satisfies the compatibility conditions

$$(\mathcal{P}_{F_1,F_2})_\# \mu_{F_2} = \mu_{F_1}, \ \ \forall F_1 \subset F_2. \tag{1.2.1}$$

Kolmogorov's existence theorem shows that conversely, given any projective family of probability measures $\mu_F : \mathcal{B}_{\boldsymbol{U}}^F \to [0,1]$, $F \in 2_0^{\boldsymbol{T}}$, there exists a unique probability measure $\mu : \mathcal{B}_{\boldsymbol{U}}^{\boldsymbol{T}} \to [0,1]$ such that

$$\mu_F = (\pi_F)_\# \mu, \ \ \forall F.$$

Thus the distribution of a random field X is uniquely determined by the distributions of the finite dimensional random vectors

$$X_F : \Omega \to \boldsymbol{U}^F, \ \ \omega \mapsto \big(X(t) \big)_{t \in F}, \ \ F \in 2_0^{\boldsymbol{T}}.$$

Definition 1.2.2. Let $(\Omega, \mathcal{S}, \mathbb{P})$ be a probability space, $\boldsymbol{T}$ a set, and $\boldsymbol{U}$ a finite dimensional real vector space. Consider stochastic processes

$$X, Y : \Omega \times \boldsymbol{T} \to \boldsymbol{U}, \ \ (t, \omega) \mapsto X_\omega(t), \ Y_\omega(t).$$

(i) The process Y is said to be a *modification* or *version* X, and we denote this $X \sim Y$, if for any $t \in \boldsymbol{T}$ there exists a negligible subset $\mathcal{N}_t$ such that

$$X_\omega(t) = Y_\omega(t), \ \ \forall \omega \in \Omega \setminus \mathcal{N}_t.$$

(ii) The processes X, Y are said to be *indistinguishable*, and we denote this $X \approx Y$, if there exists a negligible subset $\mathcal{N}$ such that

$$X_\omega(t) = Y_\omega(t), \ \ \forall t \in \boldsymbol{T}, \ \ \forall \omega \in \Omega \setminus \mathcal{N}.$$

(iii) The processes X, Y are said to be *stochastically equivalent*, and we denote this $X \sim_s Y$, if they have the same distribution, i.e., for any $F \in 2_0^{\boldsymbol{T}}$ the random vectors X_F and Y_F have the same distribution.

□

Note that $\approx, \sim, \sim_s$ are equivalence relations and

$$X \approx Y \implies X \sim Y \implies X \sim_s Y.$$

Suppose that $\boldsymbol{U}$ is equipped with an inner product with norm $\big| - \big|$ and $\boldsymbol{T}$ is a metric space. Suppose that μ is a σ-finite Borel measure on $\boldsymbol{T}$ and

$$X : \big(\Omega, \mathcal{S}, \mathbb{P} \big) \times \boldsymbol{T} \to \boldsymbol{U}.$$

In many applications we would be interested to know if the sample maps $X_\omega : \boldsymbol{T} \to \boldsymbol{U}$ have additional compatibility properties with the additional metric and measure-theoretic structures on the parameter space $\boldsymbol{T}$. In such situations measurability issues could become tricky. Let me mention two such issues

The first issue can be easily missed. It appears for example when we define random variables as a.s. limits of other random variables. Observe that

$$f, g : \big(\Omega, \mathcal{S}, \mathbb{P}\big) \to \mathbb{R},$$

are $\mathbb{P}$-a.s. equal and f is measurable, then we can conclude that g is also measurable if and only if $\mathcal{S}$ is $\mathbb{P}$-complete. To deal with this issue we will adhere to the following convention.

✍ *Unless stated otherwise, the probability spaces* $(\Omega, \mathcal{S}, \mathbb{P})$ *used throughout this book will be tacitly assumed* $\mathbb{P}$*-complete.*

To explain the second issue suppose that X is a random function defined on an open subset $\boldsymbol{T}$ of $\mathbb{R}^m$. We will have to consider quantities of the type $\sup_{t \in B} X(t)$, where B is some Borel subset of $\boldsymbol{T}$. If B is uncountable this quantity may not be measurable. This is a bit more subtle. To explain how to handle it we need a bit more terminology.

Definition 1.2.3. Let $(\boldsymbol{T}, d)$ be a metric space and X an $\boldsymbol{U}$-valued random field on $\boldsymbol{T}$.

$$X : (\Omega, \mathcal{S}, \mathbb{P}) \times \boldsymbol{T} \to \boldsymbol{U}$$

(i) The random field X is called *separable* if there exists a countable *separant*, i.e., a countable dense set $\mathcal{D} \subset \boldsymbol{T}$ and a $\mathbb{P}$-negligible subset $\mathcal{N} \subset \Omega$ such that, for any $t \in \boldsymbol{T}$, any $\varepsilon > 0$ and any $\omega \in \Omega \setminus \mathcal{N}$ we have

$$X_\omega(t) \in \boldsymbol{cl}\,\Big(\,\big\{\, X_\omega(s),\ \ s \in \mathcal{D} \cap B_\varepsilon(t)\,\big\}\,\Big).$$

(ii) The random field X said to be *stochastically continuous* if for any $t_0 \in \boldsymbol{T}$, the random variable $X(t)$ converges in probability to $X(t_0)$. More explicitly, this means that for any $t_0 \in \boldsymbol{T}$ and any $\varepsilon > 0$

$$\lim_{t \to t_0} \mathbb{P}\big[\,\big|\, X(t) - X(t_0)\,\big| > \varepsilon\,\big] = 0.$$

(iii) The random field X is called *measurable* if the map X is $\mathcal{S} \otimes \mathcal{B}_{\boldsymbol{T}}$-measurable.

(iv) The metric space $\boldsymbol{T}$ is called *convenient* if it is locally compact and separable.

□

The topology of a convenient metric space $\boldsymbol{T}$ can be defined by a complete metric whose balls are relatively compact.

Let us observe that if the random function $X : \Omega \times \boldsymbol{T} \to \mathbb{R}$ is separable, then $\boldsymbol{T}$ is separable and for any Borel subset $B \subset T$ the function

$$\Omega \ni \omega \mapsto s_B(\omega) := \sup_{t \in B} X_\omega(t) \in (-\infty, \infty]$$

is measurable since

$$s_B(\omega) = \sup_{t \in B \cap \mathcal{D}} X_\omega(t).$$

Indeed, since $\boldsymbol{T}$ is separable, the set B can be covered by countably many balls $B_{r_n}(x_n)$ and

$$\sup_B X = \sup_n \sup_{B\cap B_{r_n}(x_n)} X.$$

The separability of the random map X implies that $\sup_{B\cap B_{r_n}(x_n)} X$ is measurable for any n. We have the following result [45, Sec.II.2], [68, Sec. 4.3].

Theorem 1.2.4. *Suppose that $\boldsymbol{T}$ is a convenient space $X : (\Omega, \mathcal{S}, \mathbb{P}) \times T \to \boldsymbol{U}$ is a stochastically continuous process. Then for any σ-finite measure μ on $\mathcal{B}_{\boldsymbol{T}}$ the random field X admits a* separable *modification $Y : \Omega \times \boldsymbol{T} \to \boldsymbol{U}$ with the following additional property: there exists a $\mathbb{P}\otimes\mu$-negligible set $\mathcal{Z} \in \mathcal{S}\otimes\mathcal{B}_{\boldsymbol{T}}$ such that $\mathbb{P}\otimes\mu\big[\,\mathcal{Z}\,\big] = 0$ and the restriction of Y to $(\Omega\times\boldsymbol{T})\setminus\mathcal{Z}$ is $\mathcal{S}\otimes\mathcal{B}_{\boldsymbol{T}}$-measurable.* □

Definition 1.2.5. Let $X : \Omega \times \boldsymbol{T} \to \boldsymbol{U}$, $(\omega, t) \mapsto X_\omega(t) \in \boldsymbol{U}$ be a random field, where $\boldsymbol{U}$ is a finite dimensional Euclidean space and $\boldsymbol{T}$ is a convenient metric space. We say that X is *continuous* if for any $\omega \in \Omega$ the sample map

$$\boldsymbol{T} \ni t \mapsto X_\omega(t) \in \boldsymbol{U}$$

is continuous. The process is called a.s. *continuous* if it is indistinguishable from a continuous process.

If $\boldsymbol{T}$ is an open subset of a finite dimensional Euclidean space we can define in a similar fashion the concept of a.s. C^k random map. □

There exists sufficient conditions guaranteeing that a random map X admits a modification that is a.s. continuous. We mention here Kolmogorov's famous continuity theorem. For a proof we refer to [142, Thm. 10.1], or [148, Thm. 2.5.3].

Theorem 1.2.6 (Kolmogorov)**.** *Suppose that $\boldsymbol{T} = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ and*

$$X : \Omega \times \boldsymbol{T} \to \boldsymbol{U}$$

is a random field valued in the finite dimensional Euclidean space $\boldsymbol{U}$. If there exists $C > 0$, $p \in [1, \infty)$ and $r \in (0, 1]$ such that

$$\mathbb{E}\big[\, \|X(s) - X(t)\|^p \,\big] \leq C\big|\, s - t\,\big|^{n+pr},$$

then for any $\alpha \in (0, r]$, the random field admits a modification that is a.s. *α-Hölder continuous.* □

For more refined results of this kind we refer to [88, Chap. 11].

Example 1.2.7. (a) Suppose that $A_0, A_1, \ldots, A_n$ are independent random variables, $\boldsymbol{T} = \mathbb{R} = \boldsymbol{U}$. Define the random function

$$X : \mathbb{R} \to \mathbb{R},\quad X(t) = \sum_{k=0}^{n} A_k t^k.$$

This is an example of random polynomial. Clearly X is a.s. smooth.

(b) Suppose that A_n, B_n, $n \in \mathbb{Z}_{\geq 0}$, are independent mean zero L^2-random variables defined on the same probability space $(\Omega, \mathcal{S}, \mathbb{P})$. For simplicity we assume $\mathrm{Var}\,\big[\,A_n\,\big] = \mathrm{Var}\,\big[\,B_n\,\big] =: v_n$,

$\forall n \in \mathbb{N}$. Consider the random Fourier series

$$X : \Omega \times [0, 2\pi] \to \mathbb{R}, \;\; X_\omega(\theta) = A_0 + \sum_{n\in\mathbb{N}} \big(A_n(\omega)\cos(n\theta) + B_n(\omega)\sin(n\theta) \big). \tag{1.2.2}$$

Kolmogorov's one-series theorem shows that if

$$\sum_n \mathbb{E}\big[\, A_n^2 \,\big] + \sum_n \mathbb{E}\big[\, B_n^2 \,\big] = 2\sum_n v_n < \infty,$$

then, for any $\theta \in [0, 2\pi]$ there exists a negligible subset $\mathcal{N}_\theta \subset \Omega$ such that $\forall \omega \in \Omega \setminus \mathcal{N}_\theta$ the series

$$\sum_{n\in\mathbb{N}} \big(A_n(\omega)\cos(n\theta) + B_n(\omega)\sin(n\theta) \big)$$

is convergent. We could redefine $X(\theta)$ on $\mathcal{N}_\theta$ to be 0 and we get indeed a family of random variables on Ω parametrized by θ.

The covariance kernel of this random function is

$$\mathcal{K}(\theta, \varphi) = \sum_{n\geq 0} v_n \cos\big(n(\theta - \varphi) \big).$$

However, in our applications we would like the sample functions $\theta \mapsto X_\omega(\theta)$ to be well behaved for most ω and the above approach may prevent this from happening since the set

$$\bigcup_\theta \mathcal{N}_\theta$$

need not be negligible. For the applications we have in mind a less sophisticated ad-hoc approach will suffice. Here is a taste of this approach.

The functions $u_n(\theta) = \sin n\theta$ and $v_n(\theta) = \cos n\theta$ belong to the Banach space $C\big([0, 2\pi] \big)$ with sup-norm $\| - \|$. Moreover $\|u_n\| = \|v_n\| = 1$, $\forall n$. For the series to converge a.s. in $C\big([0, 2\pi] \big)$ it suffices that the series

$$\sum_{n\in\mathbb{N}} \big(|A_n| + |B_n| \big)$$

be a.s. convergent. For this to happen it suffices that

$$\sum_n \mathbb{P}\big[\, |A_n| > 1/n^2 \,\big] + \sum_n \mathbb{P}\big[\, |B_n| > 1/n^2 \,\big] < \infty.$$

Indeed, if the above inequality holds, then we deduce from the Borel-Cantelli lemma that

$$\mathbb{P}\big[\, |A_n| > 1/n^2 \text{ i.o.} \big] = 0 = \mathbb{P}\big[\, |B_n| > 1/n^2 \text{ i.o.} \big].$$

Thus for ω outside a negligible set we have

$$|A_n(\omega)| \leq 1/n^2 \;\text{ and }\; |B_n(\omega)| \leq 1/n^2,$$

for all but finitely many n's.

Thus, the coefficients A_n and B_n are highly concentrated near 0 for n large, thus they are very likely to be very small and we could expect that the random Fourier series describes a function that a.s. continuous. □

1.2.2. Gaussian random fields. Let $\boldsymbol{U}$ be a finite dimensional *real Euclidean* space and $\boldsymbol{T}$. A *Gaussian random field* is random field $X : (\Omega, \mathcal{S}, \mathbb{P}) \times \boldsymbol{T} \to \boldsymbol{U}$ such that, for any finite subset $F \subset \boldsymbol{T}$, the random vector $\big(X_F(t) \big)_{t \in F} \in \boldsymbol{U}^F$ is Gaussian.

Definition 1.2.8. Let $X : (\Omega, \mathcal{S}, \mathbb{P}) \times \boldsymbol{T} \to \boldsymbol{U}$ be a Gaussian field on the set $\boldsymbol{T}$.

(i) The Gaussian field X is called *centered* if $X(t)$ is a centered Gaussian vector for any $t \in \boldsymbol{T}$.

(ii) We say that X is *ample*[3] if the Gaussian vector $X(t) \in \boldsymbol{U}$ is nondegenerate for any $t \in \boldsymbol{T}$.

(iii) Given $k \in \mathbb{N}$, we say that X is *k-ample* if for any distinct points $t_1, \dots, t_k \in \boldsymbol{T}$, the Gaussian vector
$$X(t_1) \oplus \cdots \oplus X(t_k) \in \boldsymbol{U}^k$$
is nondegenerate.

(iv) The Gaussian field X is called or *∞-ample* if it is k-ample for any $k \in \mathbb{N}$.

(v) When $\boldsymbol{U} = \mathbb{R}$ we say that X is a *Gaussian function.*

□

Example 1.2.9. The random function $Z\big(2 + \sin t \big)$, Z standard normal random variable, is ample but not 2-ample since it is periodic. The random function $Z \sin t$ is not even ample.□

Suppose for simplicity that $\boldsymbol{U}$ is equipped with an inner product. For any finite subset $F \subset \boldsymbol{T}$ distribution of the random vector X_F is uniquely determined by its mean and variance.

The mean is the function
$$\boldsymbol{T} \ni t \mapsto \mathbb{E}\big[X(t) \big] \in \boldsymbol{U}.$$
The variance of X_F is a symmetric operator $\operatorname{Var}\big[X_F \big] : \boldsymbol{U}^F \to \boldsymbol{U}^F$. If $F = \{t_1, \dots, t_n\}$, then $\operatorname{Var}\big[X_F \big]$ has the block decomposition
$$\operatorname{Var}\big[X_F \big] = \big(\mathcal{K}(t_i, t_j) \big)_{1 \le i,j \le n}$$
where
$$\mathcal{K}(t_i, t_j) = \operatorname{Cov}\big[X(t_i), X(t_j) \big] \in \operatorname{Hom}(\boldsymbol{U}, \boldsymbol{U}).$$
The resulting function
$$\mathcal{K} : \boldsymbol{T} \times \boldsymbol{T} \to \operatorname{Hom}(\boldsymbol{U}, \boldsymbol{U}), \;\; (s,t) \mapsto \mathcal{K}(s,t)$$
is called the *covariance kernel* of the Gaussian field X. Recall that $2_0^{\boldsymbol{T}}$ denotes the collection of finite subsets of $\boldsymbol{T}$.

Proposition 1.2.10. *Let $\boldsymbol{U}$ be a finite dimensional Euclidean space, $\boldsymbol{T}$ a set and $\mathcal{K}$ a map*
$$\mathcal{K} : \boldsymbol{T} \times \boldsymbol{T} \to \operatorname{Hom}(\boldsymbol{U}, \boldsymbol{U}).$$
The following are equivalent.

(i) *There exists a centered Gaussian field $X : \Omega \times \boldsymbol{T} \to \boldsymbol{U}$ with covariance kernel $\mathcal{K}$.*

[3] We use the term ample since this property closely related to the ampleness condition in algebraic geometry. Many authors refer to ample fields as nondegenerate

(ii) *For any* $F \in 2_0^{\boldsymbol{T}} \subset \boldsymbol{T}$ *the operator*

$$\mathcal{K}_F : \boldsymbol{U}^F \to \boldsymbol{U}^F,$$

given by the block decomposition $\big(\mathcal{K}(f,f')\big)_{f,f' \in F}$ *is symmetric and nonnegative.*

Proof. The implication (i) ⇒ (ii) follows form the fact the variance of a centered Gaussian measure on $\boldsymbol{U}^F$ is a symmetric nonnegative operator.

Suppose that $\mathcal{K}$ satisfies (ii). For any finite subset $F \subset \boldsymbol{T}$ we denote by $\boldsymbol{\Gamma}_F$ the centered Gaussian measure on $\boldsymbol{U}^F$ with variance $\operatorname{Var}_F = \mathcal{K}_F$. The collection $\boldsymbol{\Gamma}_F$, $F \in 2_0^{\boldsymbol{T}}$ is projective in the sense of (1.2.1). Invoking Kolmogorov's existence theorem we deduce that there exists a unique probebility measure $\boldsymbol{\Gamma}_{\boldsymbol{T}}$ on $\boldsymbol{U}^{\boldsymbol{T}}$ such that, $\forall F \in 2_0^{\boldsymbol{T}}$,

$$\boldsymbol{\Gamma}_F = (\pi_F)_{\#}\boldsymbol{\Gamma}_{\boldsymbol{T}}$$

where π_F denotes the natural projection. The random field

$$\mathbf{Ev} : \boldsymbol{U}^{\boldsymbol{T}} \times \boldsymbol{T} \to \boldsymbol{U}, \;\; (u : \boldsymbol{T} \to \boldsymbol{U}) \mapsto \mathbf{Ev}_t(u) = u(t)$$

is centered Gaussian with covariance kernel $\mathcal{K}$. □

If $\boldsymbol{T}$ is a metric space and the map $t \mapsto X(t)$ is continuous in probability, then the covariance kernel $\mathcal{K}$ is a Mercer kernel in the sense of Definition 1.1.58.

Example 1.2.11. Suppose that $\boldsymbol{T} = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ and $X : \Omega \times \boldsymbol{T} \to \boldsymbol{U}$ is a centered Gaussian field such that the covariance kernel $(s,t) \mapsto \mathcal{K}(s,t)$ is Lipschitz. For $s, t \in \boldsymbol{T}$, $X(t) - X(s)$ is a Gaussian vector with variance operator

$$A_{s,t} = \operatorname{Var}\big[\, X(s) - X(t) \,\big] = \mathcal{K}(t,t) - \mathcal{K}(s,t) - \mathcal{K}(t,s) - \mathcal{K}(t,t)$$

Then, for any $k \in \mathbb{N}$ we have

$$\mathbb{E}\big[\, \|X(s) - X(t)\|^{2k} \,\big] = (2k-1)!! \operatorname{tr} A_{s,t}^k$$

Observe that

$$\|A_{s,t}\| \leq \|\mathcal{K}(t,t) - \mathcal{K}(s,t)\| + \|\mathcal{K}(t,s) - \mathcal{K}(s,s)\|.$$

since $(s,t) \mapsto \mathcal{K}(s,t)$ is locally Lipschitz we deduce that for any box $B \subset \mathscr{V}$ $\exists C = C(B) > 0$ such that

$$\|A_{s,t}\| \leq C(B)|s-t|, \;\; \forall s,t \in B.$$

Then

$$\operatorname{tr} A_{s,t}^k \leq (\dim \boldsymbol{U})^k \|A_{s,t}\|^k$$

We deduce that for $k > n$

$$\mathbb{E}\big[\, \|X(s) - X(t)\|^{2k} \big[\leq C_1|s-t|^k \leq C_2|s-t|^{n+1}, \;\; \forall s,t \in B,$$

and Kolmogorov's continuity theorem implies that the process admits a Hölder continuous modification if its covariance kernel is Lipschitz continuous. □

When X is Gaussian we can improve Kolmogorov's continuity result, Theorem 1.2.6.

Theorem 1.2.12 (Dudley)**.** *If $\boldsymbol{T}$ is a compact subset of a Euclidean space $\mathbb{R}^N$ and there exists $C > 0$, and $\alpha > 0$ such that*

$$\mathbb{E}\big[\,\|X(s) - X(t)\|^2\,\big] \leq \frac{C}{\big|\log|s-t|\,\big|^{1+\alpha}},\ \ \forall s,t \in \mathscr{V},$$

then X admits a modification that is a.s. *continuous.* □

For a proof of this result we refer to [**1**, Sec. 1.4], [**48**] or [**150**, Chap.1].

We can use the above result to produce sufficient conditions guaranteeing that the above Gaussian field is a.s. C^k, but they tend to be cumbersome; see e.g. [**1**, Thm.1.4.2]. Let us first sketch the broad contours of the argument in [**1**, Thm.1.4.2].

In order not to be distracted by heavy formalism we consider only the case $n = 1$ and $\dim \boldsymbol{U} = 1$ so that X is a Gaussian function of one real variable t.

Note that if X is to be a.s. C^1, then, as $t \to t_0$, the difference quotient $\frac{1}{t-t_0}\big(\,X(t) - X(t_0)\,\big)$ needs to converge in probability and thus in any L^p. The derivative $X'(t)$ is also a Gaussian function and we have

$$\mathbb{E}\big[\,X'(t)X(s)\,\big] = \partial_t \mathcal{K}(s,t),\ \ \mathbb{E}\big[\,X'(t)X'(s)\,\big] = \partial^2_{st}\mathcal{K}(s,t) \tag{1.2.3}$$

so $\mathcal{K}$ is at least twice differentiable in certain directions. To keeps things simple we assume that $\mathcal{K}$ is C^2. Note that for $t_0, t_1 \in \mathbb{R}$ and $h_0, h_1 \in \mathbb{R} \setminus \{0\}$

$$\frac{1}{h_0h_1}\mathbb{E}\big[\,\big(\,X(t_0+h_0) - X(t_0)\,\big)\big(\,X(t_1+h_1) - X(t_1)\,\big)\,\big]$$

$$= \frac{1}{h_0h_1}\Big(\,\mathcal{K}(t_0+h_0, t_1+h_1) - \mathcal{K}(t_0+h_0, t_1) - \mathcal{K}(t_0, t_1+h_1) + \mathcal{K}(t_0,t_1)\,\Big)$$

$$= \frac{1}{h_0h_1}\Big(\,\int_{t_1}^{t_1+h_1} \partial_{s_1}\mathcal{K}(t_0+h_0, s_1)ds_1 - \int_{t_1}^{t_1+h_1} \partial_{s_1}\mathcal{K}(t_0, s_1)ds_1\,\Big)$$

$$= \frac{1}{h_0h_1}\int_{t_1}^{t_1+h_1}\Big(\,\int_{t_0}^{t_0+h_0} \partial^2_{s_0,s_1}\mathcal{K}(s_0,s_1)ds_0\,\Big)ds_1$$

$$= \frac{1}{h_0h_1}\int_{[t_0,t_0+h_0]\times[t_1,t_1+h_1]} \partial^2_{s_0,s_1}\mathcal{K}(s_0,s_1)ds_0ds_1 =: \widehat{\mathcal{K}}\big(\,t_0,h_0;t_1,h_1\,\big).$$

The covariance kernel $\widehat{\mathcal{K}} : \big(\,\mathbb{R}\times\mathbb{R}^*\,\big)^2 \to \mathbb{R}$ extends by continuity to a Mercer kernel

$$\widehat{\mathcal{K}} : \mathfrak{X}^2 \to \mathbb{R},\ \ \mathfrak{X} := \mathbb{R}^2.$$

This defines a Gaussian field on $\mathbb{R}^2$ and, if the kernel $\widehat{\mathcal{K}}$ is locally Lipschitz, then X is a.s. C^1. This happens for example when $\mathcal{K}$ is C^3. More generally, if $\mathcal{K} \in C^{2\ell+1}$, then X is a.s. C^ℓ. A similar results holds if X depends on several Euclidean variables.

Definition 1.2.13 (Jets)**.** For any function $f \in C^\ell(\mathscr{V})$ we define its ℓ-th jet at v to be the vector

$$J_\ell f(v) = f(v) \oplus Df(v) \oplus \cdots \oplus D^\ell f(v),$$

where $D^k f(v)$ denotes the k-th order differential of f at v viewed as a symmetric k-linear form on $\boldsymbol{V}$. □

Theorem 1.2.14 (Nazarov-Sodin). *Fix $\ell \in \mathbb{N}_0$ and $\alpha \in (0,1)$. Suppose that $\mathscr{V}$ is an open subset of $\mathbb{R}^m$ and $X : \Omega \times \mathscr{V} \to \mathbb{R}$ is a centered Gaussian function with covariance kernel $\mathcal{K}$. Assume that*

$$\mathcal{K} \in C^{2\ell+2}\big(\mathcal{V} \times \mathcal{V}\big).$$

Fix a ball $B \subset \mathscr{V}$, $r < \operatorname{dist}(B, \partial\mathscr{V})$ and set

$$B_{+r} := \big\{ v \in \mathscr{V};\ \operatorname{dist}(v, B) \leq r \big\}.$$

Then the following hold.

(i) *The random function X is a.s. $C^{\ell,\alpha}$.*

(ii) *The ℓ-th jet $J_\ell X(v)$ is a Gaussian vector for any $v \in \mathscr{V}$.*

(iii) *For every closed ball $B \subset \mathscr{V}$ and for every compact set $S \subset \mathscr{V}$ that contains B in interior, there exists a constant $C = C(\operatorname{vol}[B], r, m, \ell, \alpha) > 0$ such that*

$$\mathbb{E}\big[\, \|X\|_{C^{\ell,\alpha}(B)} \,\big] \leq C \big\| \mathcal{K} \big\|^{1/2}_{C^{2\ell+2}(B_{+r} \times B_{+r})}, \tag{1.2.4}$$

where $C^{k,\alpha}$ denotes the spaces of functions that are k times differentiable and the k-th differential is Hölder continuous with exponent α.

□

For a proof we refer to [107, Appendix A.9].

Definition 1.2.15. Fix $\ell \in \mathbb{N}$ and $0 \leq k \leq \ell$. Suppose that $\mathcal{V}$ is an open subset of $\mathbb{R}^m$ and $X : \Omega \times \mathcal{V} \to \mathbb{R}$ is a centered Gaussian function that is a.s. C^ℓ. The random functiion is said to be J_k-ample if, for any $v \in \mathcal{V}$ the Gaussian vector $J_k X(v)$ is nondegenerate. □

Example 1.2.16 (Random linear combinations of maps)**.** Suppose that $\boldsymbol{T} = \mathbb{R}$ and $(X_k)_{0 \leq k \leq n}$ are independent standard normal random variables. Then

$$X(t) = \sum_{k=0}^{n} X_k t^k$$

is a centered Gaussian random function. It is a random polynomial of degree $\leq n$ so it is a.s. continuous. Its covariance kernel is the function

$$\mathcal{K} : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ \ \mathcal{K}(s,t) = \sum_{k=0}^{n} (st)^k.$$

(b) Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean space, $\boldsymbol{T}$ is a metric space and

$$f_1, \dots, f_N : \boldsymbol{T} \to \boldsymbol{U}$$

are continuous functions satisfying the *geometric ampleness condition*

$$\forall t \in \boldsymbol{T},\ \ \operatorname{span}\big\{ f_1(t), \dots, f_N(t) \big\} = \boldsymbol{U}.$$

If $X_1, \dots, X_N$ are independent standard normal random variables, then

$$X(t) = \sum_{k=1}^{N} X_k f_k(t)$$

is an ample continuous Gaussian field. □

Example 1.2.17 (Random trigonometric polynomials with given Netwon polyhedron)**.** Denote by $\mathbb{T}^m$ the m-dimensional torus $\mathbb{T}^m = \mathbb{R}^m/(2\pi\mathbb{Z})^m$. For $\vec{\ell} \in \mathbb{Z}^m$ and $\vec{\theta} \in \mathbb{R}^m$ we set

$$\langle \vec{\ell}, \vec{\theta} \rangle := \ell_1\theta_1 + \cdots + \ell_m\theta_m,$$

and

$$|\,\vec{\ell}\,| := \max_{i\leq k\leq m} |\ell_k|.$$

Fix $N \in \mathbb{N}$ and a convex polyhedron $P \subset \mathbb{R}^m$ satisfying the following properties.

- The vertices of P are lattice points, i.e., points in $\mathbb{Z}^m$.
- The origin is contained in the interior of P, $0 \in \boldsymbol{int}\, P$.
- The polyhedron is symmetric with respect to the origin, i.e., $\boldsymbol{x} \in P \Longleftrightarrow -\boldsymbol{x} \in P$.

Denote by $\prec$ the lexicographic order on $\mathbb{R}^m$ where $\boldsymbol{x} \prec \boldsymbol{y}$ iff there exists j such that $x_j < y_j$ and $x_i = y_i$, $\forall i < j$. The lexicographic order is a total (linear) order and $\boldsymbol{x} \prec \boldsymbol{y} \Longleftrightarrow -\boldsymbol{y} \prec -\boldsymbol{x}$. Fix independent standard normal random variables

$$A_{\vec{\ell}},\ B_{\vec{k}},\ \ \vec{\ell} \succeq 0,\ \ \vec{k} \succ 0$$

and set

$$Z_{\vec{\ell}} = \begin{cases} A_0, & \vec{\ell} = 0,\\ \frac{1}{\sqrt{2}}\big(A_{\vec{\ell}} - \boldsymbol{i}B_{\vec{\ell}}\big), & \vec{\ell} \succ 0,\\ \frac{1}{\sqrt{2}}\big(A_{-\vec{\ell}} + \boldsymbol{i}B_{-\vec{\ell}}\big), & \vec{\ell} \prec 0.\end{cases}$$

We denote by P_N the dilated polygon $P_N = N \cdot P$. We have a random trigonometric polynomial

$$X_N(\vec{\theta}) = \sum_{\vec{\ell}\in P_N} Z_{\vec{\ell}} e^{\boldsymbol{i}\langle \vec{\ell}, \vec{\theta}\rangle} = A_0 + \sum_{\substack{\vec{\ell}\in P_N\\ \vec{\ell}\succ 0}} \sqrt{2}\big(A_{\vec{\ell}}\cos\langle \vec{\ell}, \vec{\theta}\rangle + B_{\vec{\ell}}\sin\langle \vec{\ell}, \vec{\theta}\rangle\big). \tag{1.2.5}$$

The *Newton polyhedron* of X_N is a.s. P_N.

The random trigonometric polynomial $X_N(\vec{\theta})$ is a centered Gaussian function with covariance function

$$\mathcal{K}\big(\vec{\theta}, \vec{\varphi}\big) = \sum_{\vec{\ell}\in P_N} \cos\langle \vec{\ell}, \vec{\theta} - \vec{\varphi}\rangle = \sum_{\vec{\ell}\in P_N} e^{\boldsymbol{i}\langle \vec{\ell}, \vec{\theta}-\vec{\varphi}\rangle}.$$

If we set $\vec{\tau} := \vec{\theta} - \vec{\varphi}$ we deduce

$$\mathcal{K}\big(\vec{\theta}, \vec{\varphi}\big) = \underbrace{\sum_{\vec{\ell}\in P_N} e^{\boldsymbol{i}\langle \vec{\ell}, \vec{\tau}\rangle}}_{=:S_N(\vec{\tau})} \tag{1.2.6}$$

Note that $\vec{\tau} \mapsto S_N(\vec{\tau})$ is an even function. For any multi-index $\alpha \in \mathbb{Z}^m_{\geq 0}$, and any $\boldsymbol{x} \in \mathbb{R}^n$ we set

$$|\alpha| := \sum_{j=1}^m \alpha_j,\ \ \boldsymbol{x}^\alpha := \prod_{k=1}^m x_k^{\alpha_k}.$$

We have

$$\partial_\tau^\alpha S_N(0) = \sum_{\vec{\ell}\in P^N} \boldsymbol{i}^{|\alpha|}\vec{\ell}^{\alpha}. \tag{1.2.7}$$

Using Riemann sums one can show that

$$\lim_{N\to\infty}\frac{1}{N^{m+|\alpha|}}S_N(0)=\boldsymbol{i}^{|\alpha|}\underbrace{\int_P \boldsymbol{x}^\alpha\,d\boldsymbol{x}}_{=:\mu_\alpha[\,P\,]}.$$

We can be a bit more precise. The results in [**26**] show that

$$\partial_\tau^\alpha S_N(0)=\boldsymbol{i}^{|\alpha|}N^{m+|\alpha|}\mu_\alpha\big[\,P\,\big]\big(\,1+O(1/N)\,\big)\ \text{ as } N\to\infty. \tag{1.2.8}$$

We deduce that

$$\mathrm{Var}\big[\,X_N(\vec{\theta})\,\big]=\mathcal{K}\big(\,\vec{\theta},\vec{\theta}\,\big)=S_N(0)=N^m\,\mathrm{vol}\big[\,P\,\big]\big(\,1+O(1/N)\,\big).$$

Hence $X_N(\vec{\theta})$ is nondegenerate for any $\vec{\theta}$ if N is sufficiently large. In other words X_N is ample for $N\gg 0$.

The random trigonometric polynomial X_N is C^∞. If $\boldsymbol{e}_1,\dots,\boldsymbol{e}_m$ denotes the canonical basis of $\mathbb{R}^m$, then we have a.s.

$$\partial_{\theta_i}X_N(\vec{\theta})=\lim_{h\to 0}\frac{1}{h}\big(\,X_N(\vec{\theta}+h\boldsymbol{e}_i)-X_N(\vec{\theta})\,\big) \tag{1.2.9}$$

The variables in the right-hand-side of the above equality are Gaussian. Hence the limit is also Gaussian and the convergence to the limit holds in any L^p, $p\in[1,\infty)$. This proves that the gradient $\nabla X_N(\vec{\theta})$ is an $\mathbb{R}^m$-valued Gaussian field. Then same argument shows that $X_N,\nabla X_N$ are jointly Gaussian.

If $|\alpha|$ is odd, then $\partial_\tau^\alpha S_N(0)=0$ since P is symmetric with respect to the origin. The equality (1.2.9) implies that

$$\mathrm{Cov}\big[\,\partial_{\theta_i}X_N(\vec{\theta}),X_N(\vec{\theta})\,\big]=\partial_{\tau_i}S_N(0)=0$$

so that $\mathrm{Cov}\big[\,X_n(\vec{\theta}),\nabla X_N(\vec{\theta})\,\big]=0$. Thus $X_N(\vec{\theta})$ and $\nabla X_N(\vec{\theta})$ are independent for any $\vec{\theta}$.

The covariance kernel of $\nabla X_N(\vec{\theta})$ and $\nabla X_N(\vec{\varphi})$ is given by the linear operator

$$\mathcal{K}^\nabla(\vec{\theta},\vec{\varphi}):\mathbb{R}^m\to\mathbb{R}^m$$

described by the $m\times m$ matrix

$$\mathcal{K}^\nabla(\vec{\theta},\vec{\varphi})_{ij}=\mathbb{E}\big[\,\partial_{\theta_i}X_N(\vec{\theta})\partial_{\varphi_j}X_N(\vec{\varphi})\,\big]=\partial_{\theta_i}\partial_{\varphi_j}\mathcal{K}(\vec{\theta},\vec{\varphi}\,).$$

The variance operator $\mathrm{Var}\big[\,\nabla X_N(\vec{\theta})\,\big]$ is described by the symmetric $m\times m$ with entries

$$\partial_{\theta_i}\partial_{\varphi_j}\mathcal{K}(\vec{\theta},\vec{\varphi}\,)_{\vec{\theta}=\vec{\phi}}=-\partial^2_{\tau_i\tau_j}S_N(0)\sim\mu_{ij}\big[\,P\,\big]N^{m+2}\ \text{ as } N\to\infty, \tag{1.2.10}$$

where

$$\mu_{ij}\big[\,P\,\big]=\int_P x_ix_jd\boldsymbol{x}.$$

The matrix of moments

$$M(P):=\big(\,\mu_{ij}\big[\,P\,\big]\,\big)_{1\le i,j\le m} \tag{1.2.11}$$

is the Gramian matrix of the functions $\ell_i:P\to\mathbb{R}$, $\ell_i(x_1,\dots,x_m)=x_i$, $i=1,\dots,m$, with respect to the inner product in $L^2(P,\boldsymbol{\lambda})$. These functions are linearly independent since the interior of P is nonempty. Thus the matrix $M(P)$ of moments is invertible. From the asymptotic equality equality

$$\mathrm{Var}\big[\,\nabla X_n(\vec{\theta})\,\big]\sim N^{m+2}M(P)\ \ N\to\infty \tag{1.2.12}$$

the Gaussian map ∇X_N is ample for all N sufficiently large. In particular, this also shows that X_N is J_1-ample if N is large.

A similar argument shows that for any $k \in \mathbb{N}$, X_N is J_k-ample if N is sufficiently large. □

1.2.3. Gaussian fields and Gaussian measures. The concepts of Gaussian fields and Gaussian measures are intimately related. In this subsection we describe mainly through examples different facets of this relationship.

Example 1.2.18. Suppose that M is a compact metric space.

Suppose that Γ is a Gaussian measure on $\boldsymbol{F} = C(M)$. We obtain a probability space $(\boldsymbol{F}, \mathcal{B}_{\boldsymbol{F}}, \Gamma)$ and a random Gaussian function

$$E^{\Gamma} : \boldsymbol{F} \times M \to \mathbb{R}, \;\; (f, x) \mapsto E^{\Gamma}_f(x) = \mathbf{Ev}_x(f) = f(x).$$

This is a centered Gaussian random Gaussian function that is, tautologically, continuous, i.e., for any $f \in \boldsymbol{F}$ the sample map $x \mapsto f(x)$ is continuous. Since the map E^{Γ} is continuous it is also Borel measurable so the associated random function is measurable in the sense of Definition 1.2.3.

The covariance kernel of this random function coincides with the covariance kernel of the Gaussian measure Γ constructed in Example 1.1.59. In particular, it is a Mercer kernel

$$K^{\Gamma} : M \times M \to \mathbb{R} = \mathbb{E}_{\Gamma}\big[\, \mathbf{Ev}_x \, \mathbf{Ev}_y \,\big].$$

Let us point out that the Gaussian measure Γ can also be viewed as a Gaussian random function on $\boldsymbol{F}^*$

$$\Phi^{\Gamma} : \boldsymbol{F} \times \boldsymbol{F}^* \to \mathbb{R}, \;\; \Phi_f(\xi) = \langle \xi, f \rangle.$$

There is a natural map $M \to \boldsymbol{F}^*$, $\mathbf{Ev} : M \to \boldsymbol{F}^*$, $x \mapsto \mathbf{Ev}_x$. The random function E^{Γ} is the pullback of Φ^{Γ} by $\mathbf{Ev}$,

$$E^{\Gamma}(x) = \Phi^{\Gamma}\big(\, \mathbf{Ev}_x \,\big).$$

Conversely, suppose that

$$\Psi : (\Omega, \mathcal{S}, \mathbb{P}) \times M \to \mathbb{R}, \;\; (\omega, t) \mapsto \Psi_{\omega}(x)$$

is a centered Gaussian random function that is a.s. continuous. Thus there exists a negligible subset $\mathcal{N} \in \mathcal{S}$ such that, $\forall \omega \in \Omega \setminus \mathcal{N}$ the function $M \ni x \mapsto \Psi_{\omega}(x)$ is continuous. Modify Ψ so that $\Psi_{\omega} : M \to \mathbb{R}$ is identically zero for $\omega \in \mathcal{N}$. We obtain a measurable map

$$\Phi : \Omega \to \mathbb{R}^M, \;\; \Omega \ni \omega \mapsto X_{\omega} \in \mathbb{R}^{\boldsymbol{T}},$$

whose image is contained in $C(M)$. Since the Borel sigma-algebra of $\boldsymbol{F} = C(M)$ is the restriction of the product sigma-algebra $\mathcal{B}_{\mathbb{R}}^M$ we deduce that Ψ defines a measurable map

$$\Psi : (\Omega, \mathcal{S}, \mathbb{P}) \to \big(\, \boldsymbol{F}, \mathcal{B}_{\boldsymbol{F}} \,\big), \;\; \omega \mapsto \Psi_{\omega}.$$

We will show that $\mu = \Psi_{\#}\mathbb{P}$ is a Gaussian measure on $\boldsymbol{F}$.

Denote by $\mathcal{L}$ the subspace of $\boldsymbol{F}^*$ spanned by the evaluation maps $\mathbf{Ev}_x$, $x \in M$. The random function Ψ is Gaussian so, for any finite subset $\{x_1, \dots, x_n\} \subset M$ and any $c_1, \dots, c_n \in \mathbb{R}$, the random variable $c_1\Psi(t_1) + \cdots + c_n\Psi(t_n)$ is Gaussian. In other words, if

$$\xi = \sum_{k=1}^{n} c_k \, \mathbf{Ev}_{t_k} \in \mathcal{L},$$

then $\langle \xi, \Psi \rangle$ is Gaussian, i.e. $\xi_{\#}\mu$ is Gaussian. Since $\sigma(\mathcal{L}) = \sigma(\boldsymbol{F}^*) = \mathcal{B}_{\boldsymbol{F}}$, we deduce from Corollary 1.1.49 that μ is centered Gaussian. By construction, the processes $\big(\Psi(x)\big)_{x \in M}$ and $\big(E^\Gamma(x)\big)_{x \in M}$ have the same distribution.

Arguing in a similar fashion one can show that if $\boldsymbol{U}$ is a finite dimensional Euclidean space, then any a.s. continuous centered Gaussian field $\Psi : M \to \boldsymbol{U}$ determines a centered Gaussian measure Γ on the Banach space $C(M, \boldsymbol{U})$ and conversely, any Gaussian measure on this Banach space is determined in this fashion.

In this case for any $x \in M$ and $\boldsymbol{u}^* \in \boldsymbol{U}$ we have an evaluation map

$$\mathbf{Ev}_{x|u^*} : C(M, \boldsymbol{U}) \to \mathbb{R}, \;\; f \mapsto \big(f(x), \boldsymbol{u}^*\big),$$

where $(-,-)$ denotes the inner product on $\boldsymbol{U}$. The sigma-algebra generated by these continuous functionals generates the Borel sigma algebra of $C(\boldsymbol{T}, \boldsymbol{U})$. For any $x_0, x_1 \in M$, the covariance operator

$$\mathcal{K}_\Psi(x_1, x_0) : \boldsymbol{U} \to \boldsymbol{U}$$

is uniquely determined by the equality

$$\big(\boldsymbol{u}_1, \mathcal{K}_\Psi(x_1, x_0)\boldsymbol{u}_0\big) = \mathbb{E}_\Gamma\big[\, \mathbf{Ev}_{x_1|\boldsymbol{u}_1} \mathbf{Ev}_{x_0|\boldsymbol{u}_0} \,\big] = \mathrm{Cov}_\Gamma\big[\, \mathbf{Ev}_{x_1|\boldsymbol{u}_1}, \mathbf{Ev}_{x_0|\boldsymbol{u}_0} \,\big],$$

$\forall \boldsymbol{u}_0, \boldsymbol{u}_1 \in \boldsymbol{U}$. □

Example 1.2.19. Let M be smooth, compact connected m-dimensional submanifold of a Euclidean space $\boldsymbol{U}$. Denote by g the induced metric on M and by vol_g the volume measure determined by g. Set $\boldsymbol{F} = C^0(M)$. We can use the metric to define a sup-like norm on $C^1(M)$ and we denote by $\boldsymbol{F}_1$ the resulting Banach space.

The inclusion $\boldsymbol{F}_1 \to \boldsymbol{F}$ is continuous. Suppose that Γ is a Gaussian measure on $\boldsymbol{F}_1$. We obtain as before a Gaussian process

$$E^\Gamma : \boldsymbol{F}_1 \times M \to \mathbb{R}, \;\; (f, x) \mapsto \mathbf{Ev}_x(f)$$

It is tautologically C^1 and its covariance kernel coincides with the covariance kernel of the Gaussian measure Γ.

Conversely, any centered Gaussian C^1-field $\Psi : \Omega \times \mathcal{O} \to \mathbb{R}$ determines as in Example 1.2.18 a Gaussian measure Γ on $\boldsymbol{F}_1 = C^1(M)$ such that the processes $\big(\Psi_x\big)_{x \in M}$ and $\big(E^\Gamma(x)\big)_{x \in M}$ have the same distribution.

Their common distribution is determined by the covariance kernel $\mathcal{K}$ of Ψ,

$$\mathcal{K} : M \times M \to \mathbb{R}, \;\; \mathcal{K}(x_0, x_1) = \mathbb{E}\big[\, \Psi(x_0)\Psi(x_1) \,\big].$$

Fix two tangent vectors $v_i \in T_{x_i}M \subset \boldsymbol{U}$, $i = 0, 1$. Let us observe that the directional derivatives $\partial_{v_0}\Psi(x_0)$ and $\partial_{v_1}\Psi(x_1)$ are jointly Gaussian.

To see this choose smooth paths $\gamma_i : (-1, 1) \to M$, $i = 0, 1$, such that

$$\gamma_i(0) = x_i, \;\; \dot{\gamma}_i(0) = v_i.$$

Then

$$\left[\begin{array}{c} \partial_{v_0}\Psi(x_0) \\ \partial_{v_1}\Psi(x_1) \end{array}\right] = \lim_{h \to 0} \frac{1}{h} \left[\begin{array}{c} \Psi\big(\gamma_0(h)\big) - \Psi(x_0) \\ \Psi\big(\gamma_1(h)\big) - \Psi(x_1) \end{array}\right].$$

The random vectors on the right-hand-side are Gaussian and converge pointwisely the left-hand-side. We deduce from Proposition 1.1.31 that this convergence is in any L^p, $1 \leq p < \infty$ and the limit is also a Gaussian vector. Moreover

$$\mathbb{E}\big[\,\partial_{v_i}\Psi(x_0)\Psi(x_1)\,\big] = \lim_{h\to 0}\frac{1}{h}\Big(\,\mathbb{E}\big[\,\Psi(\gamma_0(h))\Psi(x_1)\,\big] - \mathbb{E}\big[\,\Psi(x_0)\Psi(x_1)\,\big]\,\Big)$$

$$= \lim_{h\to 0}\big(\,\mathcal{K}(\gamma_0(h),x_1) - \mathcal{K}(x_0,x_1)\,\big) = \partial_{v_0}\mathcal{K}(x_0,x_1).$$

Arguing similarly we deduce

$$\mathbb{E}\big[\,\partial_{v_0}\Psi(x_0)\partial_{v_1}\Psi(x_1)\,\big] = \partial_{v_0}\partial_{v_1}\mathcal{K}(x_0,x_1). \tag{1.2.13}$$

□

Example 1.2.20. Consider a random Taylor series of the form

$$X : \Omega\times[-1,1]\to\mathbb{R},\;\; X(t) = A_0c_0 + \sum_{n\geq 2}A_nc_nf_n(t), \tag{1.2.14}$$

where the coefficients A_n are independent centered Gaussians, $f_n(t) = t^n$, and the *positive* real numbers $c_0, c_2, \dots,$ satisfy

$$\sum_{n\geq 2}c_n < \infty. \tag{1.2.15}$$

Note that

$$\sup_{t\in[-1,1]}|f_n(t)| \leq 1,\;\;\forall n.$$

We deduce from Proposition 1.1.57 that the random series (1.2.14) converges a.s. in the Banach space $\boldsymbol{F}_0 = C\big(\,[-1,1]\,\big)$ and defines a Gaussian measure Γ_0 on this space. In particular X is a continous Gaussian random function.

The Stone-Weierstrass theorem shows that

$$V = \operatorname{span}\big\{1, f_2(t), f_3(t), \dots\,\big\}$$

is dense in this Banach space. Proposition 1.1.57 implies that the induced Gaussian measure is nondegenerate. Proposition 1.2.23 implies that X is k-ample, for any $k\in\mathbb{N}$.

Suppose now that the sequence (c_n) satisfies the more stringent requirement

$$\sum_{n\geq 2}nc_n < \infty. \tag{1.2.16}$$

We have

$$\sup_{t\in[-1,1]}\big|\,f_n'(t)\,\big| \leq n,\;\;\forall n,$$

we deduce from Proposition 1.1.57 that the random series (1.2.14) converges a.s. in the Banach space $\boldsymbol{F}_1 := C^1\big(\,[-1,1]\,\big)$ and defines a Gaussian Γ_1 measure on this space. This Gaussian measure is degenerate since $\mathbb{P}\big[\,f'(0)\neq 0\,\big] = 0$.

If we denote by $\mathcal{L}$ the span in $\boldsymbol{F}^*$ of the linear functionals $\mathbf{Ev}_t$, $t\in(-1,1)$, then $\sigma(\mathcal{L}) = \mathcal{B}_{\boldsymbol{F}}$. Moreover, any $\xi\in\mathcal{L}\setminus\{0\}$ is a nondegenerate Gaussian random variable since Γ_0 is nondegenerate. On the other hand, the linear functional $\xi_0\in\boldsymbol{F}_1^*$ given by $\xi_0(f) = f'(0)$ is degenerate. □

Theorem 1.2.21. *Fix $\ell \in \mathbb{N}_0$ and $\alpha \in (0,1)$. Suppose that $\mathcal{V}$ is an open subset of $\mathbb{R}^m$ and $X : \Omega \times \mathcal{V} \to \mathbb{R}$ is a centered Gaussian function with covariance kernel $\mathcal{K}$. Assume that*

$$\mathcal{K} \in C^{2\ell+2}\big(\mathcal{V} \times \mathcal{V}\big).$$

Then X is a.s. $C^{\ell,\alpha}$ and for every $p \in [1,\infty)$ and every box $B \subset \mathcal{V}$ there exists a constant $C_p = C_p(B, \mathcal{V}, \ell, \alpha) > 0$ such that

$$\mathbb{E}\big[\, \|X\|^p_{C^{\ell,\alpha}(B)} \,\big] \le C_p \big\| \,\mathcal{K}\, \big\|^{\frac{p+1}{2}}_{C^{2\ell+2}(B\times B)}, \tag{1.2.17}$$

where $C^{k,\alpha}$ denotes the spaces of functions that are k times differentiable and the k-th differential is Hölder continuous with exponent α.

Proof. For simplicity, we denote by $\|-\|$ the norm $\|-\|_{C^{\ell,\alpha}(B)}$ and we set

$$Z(K) := \big\| \,\mathcal{K}\, \big\|_{C^{2\ell+2}(B\times B)}.$$

According to (1.2.4) there exists a constant $C = C(B, \mathcal{V}, \ell, \alpha) > 0$ such that

$$\mathbb{E}\big[\, \|X\| \,\big] \le C Z(K)^{1/2}.$$

From Markov's inequality we deduce that

$$\mathbb{P}\big[\, \|X\| > t \,\big) \le \frac{C Z(K)^{1/2}}{t}.$$

If we choose $r_0 := 4 C Z(K)^{1/2}$, then we deduce that

$$\mathbb{P}\big[\, \|X\| > r_0 \,\big] < \frac{1}{4}.$$

The restriction $X|_B$ induces a Gaussian measure $\boldsymbol{\Gamma}$ on $C^{\ell,\alpha}(B)$. Fernique's inequality (1.1.31) applied to $\boldsymbol{\Gamma}$ shows that there exists a universal constant $\beta > 0$ such that

$$\mathbb{P}\big[\, \|X\| > r \,\big) \le r_0 e^{-\frac{\beta r^2}{r_0^2}} = r_0 e^{-Ar^2}, \;\; A = \frac{\beta^2}{r_0^2}.$$

Then

$$\mathbb{E}\big[\, \|X\|^p \,\big] = p \int_0^\infty r^{p-1} \mathbb{P}\big[\, \|X\| > r \,\big) dr \le p r_0 \int_0^\infty r^{p-1} e^{-Ar^2} dr$$

$(s = Ar^2,\ r = \sqrt{\frac{s}{A}}\ )$

$$= \frac{p r_0}{2A^{p/2}} \underbrace{\int_0^\infty s^{p/2-1} e^{-s} ds}_{=\Gamma(p/2)} = C_p r_0^{p+1}.$$

□

Definition 1.2.22. Suppose that M is a compact metric space, $\boldsymbol{U}$ a finite dimensional Euclidean space and $\Psi : \Omega \times M \to \boldsymbol{U}$ is a continuous Gaussian field. We say that Ψ is *strongly nondegenerate* if the induced Gaussian measure $\boldsymbol{\Gamma}_\Psi$ on the Banach space $C^0(M, \boldsymbol{U})$ is nondegenerate. □

Proposition 1.2.23. *Suppose that $\Psi : \Omega \times M \to \boldsymbol{U}$ is a strongly nondegenerate continuous Gaussian field. Then Φ is ∞-ample, i.e., it is k-ample for any $k \in \mathbb{N}$.*

Proof. Let $x_1, \dots, x_k$ be k distinct points in M and $\mathcal{O} \in \boldsymbol{U}^k$ an open set. The map

$$\mathbf{Ev}_{x_1,\dots,x_k} : C^0(M, \boldsymbol{U}) \to \boldsymbol{U}^k, \;\; F \mapsto \big(F(x_1), \dots, F(x_k) \big)$$

is continuous so $\widehat{\mathcal{O}} = \mathbf{Ev}^{-1}_{x_1,\dots,x_k}\big(\mathcal{O}\big)$ is an open subset of $C^0(M, \boldsymbol{U})$. If we denote by $\boldsymbol{\Gamma}_\Psi$ the Gaussian measure on $C^0(M, \boldsymbol{U})$ induced by Φ. Then

$$\mathbb{P}\big[\big(\Psi(x_1, \dots, \Psi(x_k) \big) \in \mathcal{O} \big] = \boldsymbol{\Gamma}_\Psi\big[\widehat{\mathcal{O}} \big] > 0$$

since $\boldsymbol{\Gamma}_\Psi$ is nondegenerate. □

1.2.4. Random series. Historically, the first random functions were constructed as random Fourier series or random Taylor series, [**78**]. For Gaussian functions this not just a peculiar way of constructing them. It is a feature of this class of random functions as most of them have a description as sums of random series. More precisely we have the following result, [**21**, Thm. 3.5.1].

Theorem 1.2.24. *Sppose that Γ is a centered Gaussian measure on a separable Fréchet space $\boldsymbol{X}$ with Cameron-Martin space H_Γ. Denote by $\boldsymbol{X}^*_\Gamma$ the closure of $\boldsymbol{X}^*$ in $L^2(\boldsymbol{X}, \Gamma)$. The map $T^*_\Gamma : \boldsymbol{X}^*_\Gamma \to H_\Gamma$ in (1.1.34) is a surjective isometry. For any $h \in H_\Gamma$ we set*

$$\widehat{h} := \big(T^*_\Gamma \big)^{-1} h \in \boldsymbol{X}^*_\Gamma \subset L^2(\boldsymbol{X}, \Gamma),$$

Fix a complete orthonormal system $(h_n)_{n\in\mathbb{N}}$ in H_Γ. Then there exists a Γ-negligible subset $\mathcal{N} \subset \boldsymbol{X}$ such that for any $\boldsymbol{x} \in \boldsymbol{X} \setminus \mathcal{N}$

$$\boldsymbol{x} = \sum_{n\in\mathbb{N}} \widehat{h}_n(\boldsymbol{x}) h_n,$$

where the above convergence is in the topology of $\boldsymbol{X}$. □

The goal of this subsection is elaborate on this result and see how it looks in concrete situations.

Let $\boldsymbol{T}$ be a compact metric space. The distribution of the (centered) Gaussian function on $\boldsymbol{T}$

$$\Psi : \Omega \times \boldsymbol{T} \to \mathbb{R}, \;\; (\omega, t) \mapsto \Psi_\omega(t)$$

is uniquely determined by its covariance kernel

$$K : \boldsymbol{T} \times \boldsymbol{T} \to \mathbb{R}, \;\; K(x, y) = \mathbb{E}\big[\Psi(x)\Psi(y) \big].$$

Note that K satisfies the following conditions.

(i) $K(s, t) = K(t, s)$, $\forall s, t \in \boldsymbol{T}$.

(ii) For any $t_1, \dots, t_n \in M$ the symmetric matrix $\big(K(t_i, t_j) \big)_{1\leq i,j\leq n}$ is nonnegative definite.

Conversely, Kolmogorov's existence theorem shows that any function $K : \boldsymbol{T} \times \boldsymbol{T} \to \mathbb{R}$ satisfying (i) and (ii) is the covariance kernel of a centered Gaussian function Ψ on M.

Proposition 1.2.25. *The following are equivalent.*

(i) *The Gaussian random function Ψ is stochastically continuous; see Definition 1.2.3.*

(ii) *The covariance kernel is continuous.*

Proof. (i) $\Rightarrow$ (ii). If $(s_n, t_n) \to (s, t)$, then $\Psi(s_n) \to \Psi(s)$ and $\Psi(t_n) \to \Psi(t)$ in probability, and thus also in L^2. We deduce that

$$\lim_{n\to\infty} K(s_n, t_n) = \lim_{n\to\infty} \mathbb{E}\big[\, \Psi(s_n)\Psi(t_n)\,\big] = \mathbb{E}\big[\, \Psi(s)\Psi(t)\,\big] = K(s,t)$$

(ii) $\Rightarrow$ (ii) Note that

$$\mathbb{E}\Big[\,\big(\,\Psi(t_n) - \Psi(t)\,\big)^2\,\Big] = K(t_n, t_n) - 2K(t_n, t) + K(t,t) \to 0 \;\text{ as } n \to \infty.$$

$\square$

Thus, if Ψ is stochastically continuous its covariance kernel is a Mercer kernel. If K satisfies additional conditions such as the one in Dudley's Theorem 1.2.12, then Ψ admits a modification that is continuous. In particular in this case K is continuous and thus it is a Mercer kernel.

Let us point out that not every stochastically continuous Gaussian function admits a continuous modification; see e.g. [**1**, Cor. 1.5.5] or [**17**]. However, if we know a priori that Ψ is a continuous Gaussian function, then this Gaussian function can be described as the sum of a certain random series of functions. Here are the details.

Denote by $\boldsymbol{F}$ the Banach space $C(\boldsymbol{T})$ and let $\boldsymbol{F}_1 \subset \boldsymbol{F}$ be a subspace equipped with a norm that makes it into a Banach space and the inclusion $\boldsymbol{F}_1 \hookrightarrow \boldsymbol{F}$ is continuous. E.g., $\boldsymbol{F}_1$ could be $C^1(\boldsymbol{T})$ if $\boldsymbol{T}$ were a compact smooth manifold.

Suppose that

$$\Psi : \Omega \times \boldsymbol{T} \to \mathbb{R}, \;\; (\omega, x) \mapsto \Psi_\omega(x)$$

is a centered Gaussian function such that, $\forall \omega \in \Omega$ the functions $\Psi_\omega(-)$ belongs to $\boldsymbol{F}_1$. Denote by K its covariance kernel. In particular, K is a Mercer kernel.

Arguing as in Example 1.2.19 we deduce that Ψ defines a Gaussian measure Γ on $\boldsymbol{F}_1$ whose covariance kernel coincides with the covariance kernel of Ψ. Moreover, for any $s \in \boldsymbol{T}$ the function $K_s : \boldsymbol{T} \to \mathbb{R}$, $K_s(t) = K(s,t)$ belongs to $\boldsymbol{F}_1$.

Let H_Γ denote the Cameron-Martin space of Γ. Recall that $H_\Gamma \subset \boldsymbol{F}_1$. As explained in Appendix B.5, H_Γ is the closure of the vector space

$$\operatorname{span}\big\{ K_s;\;\; s \in \boldsymbol{T}\,\big\}$$

with respect to the inner product

$$\big(\,K_s, K_t\,\big) = K(s,t), \;\; \forall s, t \in \boldsymbol{T}.$$

Equivalently, if we denote by H_Ψ the closure in $L^2(\Omega, \mathcal{S}, \mathbb{P})$ of the span of the random variables $\big(\,\Psi(t)\,\big)_{t\in\boldsymbol{T}}$, then the map

$$H_\Psi \ni \Psi(t) \mapsto K_t \in H_\Gamma$$

induces a Hilbert space isomorphism $E_\psi : H_\Psi \to H_\Gamma$; see Example B.5.5. For each $h \in H_\Gamma$ we denote $\widehat{h}$ the unique random variable in H_Ψ that corresponds to h under this isomorphism. More formally, $\widehat{h} = E_\Psi^{-1}(h)$. The space H_Γ is separable. If $(h_n)_{n\in\mathbb{N}}$ is a complete orthonormal basis of H_Γ, then $\big(\,\widehat{h}_n\,\big)$ is a sequence of independent standard normal random variables in $L^2(\Omega, \mathcal{S}, \mathbb{P})$.

We then have the following nontrivial probabilistic result.

Theorem 1.2.26 (Karhune-Loève expansion)**.** *Suppose that* $(h_n)_{n\in\mathbb{N}}$ *is an orthonormal basis of* H_Γ*. Then the random series of functions in* $\boldsymbol{F}_1$

$$S(\omega) = \sum_{n\geq 1} \widehat{h}_n(\omega) h_n$$

converges a.s. in the norm of $\boldsymbol{F}_1$ *and it is a.s. equal to* Ψ_ω*.* □

For a proof we refer to [**1**, Thm. 3.1.1], [**69**, Thm. 2.6.10], or [**148**, Thm. 4.1.1].

Using the covariance kernel of Ψ one can explicitly describe an orthonormal basis of H_Γ.

Fix a *diffuse* finite Borel probability measure μ on $\boldsymbol{T}$. Recall that this means that $\mu[\,U\,] > 0$ for any open subset of $\boldsymbol{T}$. Suppose that K is an arbitrary Mercer kernel on $\boldsymbol{T}$. As described in (1.1.35), the covariance kernel K defines a symmetric nonnegative definite integral operator

$$[K]_\mu : L^2(\boldsymbol{T},\mu) \to L^2(\boldsymbol{T},\mu).$$

This operator is compact, symmetric and nonnegative. Each nonzero eigenvalue is positive and has finite multiplicity. Let $(\lambda_n)_{n\geq 1}$ be these nonzero eigenvalues repeated according to their multiplicities. We choose an orthonormal system of $L^2(M,\mu)$ consisting of eigenfunctions of $[K]_\mu$ corresponding to these *nonzero* eigenvalues

$$(\psi_n)_{n\in\mathbb{N}},\ \ [K]_\mu\psi_n = \lambda_n\psi_n,\ \ \int_{\boldsymbol{T}} \psi_n(t)\psi_m(t)\mu[\,dt\,] = \delta_{mn} = \begin{cases} 1, & m=n,\\ 0, & m\neq n.\end{cases}$$

Since $[K]_\mu\big(\,L^2(\boldsymbol{T})\,\big) \subset C(\boldsymbol{T})$ we deduce that each ψ_n is continuous.

Theorem 1.2.27 (Mercer)**.** *The following hold.*

(i) *The series*

$$\sum_{n\geq 1} \lambda_n\psi_n(s)\psi_n(t)$$

converges uniformly and absolutely to $K(x,y)$*.*

(ii) *The operator* $[K]_\mu$ *is trace class and*

$$\operatorname{tr}[K]_\mu = \sum_{n\geq 1}\lambda_n = \int_{\boldsymbol{T}} K(t,t)\mu[\,dt\,].$$

(iii) *The collection* $\big(\,e_n = \sqrt{\lambda_n}\psi_n\,\big)_{n\in\mathbb{N}}$ *is a complete orthonormal basis of the RKHS space* $\mathcal{H}_K$ *determined by* K*. In particular, if* K *is the covariance kernel of a Gaussian measure* Γ *on* $\boldsymbol{F}_1$ *as in Theorem 1.2.26, then this collection is a complete orthonormal basis* $H_\Gamma = \mathcal{H}_K$*.*

(iv) *A function*

$$f(t) = \sum_{n\geq 1} c_n\psi_n(t) \in L^2(\boldsymbol{T},\mu)$$

belongs to $\mathcal{H}_K$ *iff*

$$\sum_{n\geq 1}\frac{c_n^2}{\lambda_n} < \infty.$$

□

For a proof we refer [**129**, Prop.11.8, Thm. 11.18].

Example 1.2.28. It is instructive to see how this works in a simple yet fundamental example. Let

$$K : [0,1] \times [0,1] \to \mathbb{R}, \quad K(s,t) = \min(s,t).$$

It is clearly continuous and symmetric. It is nonnegative definite since

$$K(s,t) = \big(\boldsymbol{I}_{[0,s]}, \boldsymbol{I}_{[0,t]} \big)_{L^2([0,1])}, \quad \forall s,t \in [0,1].$$

It defines a centered Gaussian process $X : (\Omega, \mathcal{S}, \mathbb{P}) \times [0,1] \to \mathbb{R}$ satisfying

$$\mathbb{E}\big[X(s) - X(t)|^2 \big] = K(t,t) - 2K(s,t) + K(s,s) = |t-s|.$$

According to Kolmogorov's continuity theorem it admits a continuous version. This version is the Brownian motion.

Let us find the eigenvalues of $[K] = [K]_{\mathrm{Leb}}$, where Leb denotes the Lebesgue measure. The equality $[K]\psi = \lambda\psi$ reads

$$\lambda\psi(t) = \int_0^t s\psi(s)ds + t\int_t^1 \psi(s)ds, \quad \forall t \in [0,1] \quad \psi \in L^2\big([0,1]\big). \tag{1.2.18}$$

If $\lambda = 0$ we deduce from Lebesgue's differentiation theorem that $\psi = 0$ a.e. so $\ker[K] = \{0\}$.

If $\lambda > 0$ we deduce from (1.2.18) that $\Psi \in C^\infty$ and $\psi(0) = 0$. Derivating (1.2.18) we deduce

$$\lambda\psi'(t) = t\psi(t) - t\psi(t) + \int_t^1 \psi(s)ds, \quad \psi(0) = 0.$$

Derivating again we deduce that

$$\lambda\psi''(t) = -\psi(t), \quad \psi(0) = 0,$$

so that

$$\psi(t) = A\sin(\mu t), \quad \mu := \frac{1}{\sqrt{\lambda}}.$$

If $\sin\mu t$ is an eigenfunction, then for any $t \in [0,1]$ we have the equality

$$\frac{1}{\mu^2}\sin\mu t = \int_0^t s\sin(\mu s)ds + t\int_t^1 \sin(\mu s)ds$$

$$= -\frac{t}{\mu}\cos(\mu t) + \frac{1}{\mu}\int_0^t \cos(\mu s)ds + \frac{t}{\mu}\cos\mu t - \frac{1}{\mu}\cos\mu = \frac{1}{\mu^2}\sin\mu t - \frac{\cos\mu}{\mu}.$$

This implies $\cos\mu = 0$, i.e., $\mu = \big(n - \frac{1}{2}\big)\pi$, $n \in \mathbb{N}$. Thus the spectrum of K is

$$\lambda_n = \frac{4}{(2n-1)^2\pi^2}, \quad n \in \mathbb{N}$$

and consists of simple eigenfunctions

$$\psi_n(t) = \sin\big((2n-1)\pi t/2\big), \quad \int_0^1 \psi_n(t)^2 dt = \frac{1}{2}.$$

The RKHS space $\mathcal{H}_K$ consists of functions $f \in C^0\big([0,1]\big) \cap L^2\big([0,1]\big)$ such that $f(0) = 0$ and

$$\sum_{n\in\mathbb{N}} n^2 \big| \big(f, \psi_n \big)_{L^2} \big|^2 < \infty.$$

We recognize above the square of the norm of the Sobolev space $L^{1,2}\big([0,1]\big)$ consisting of absolutely continuous functions with L^2 derivative. Hence

$$\mathcal{H}_K := \big\{ f \in L^{1,2}\big([0,1]\big);\;\; f(0)=0 \big\}.$$

If $(X_n)_{n\in\mathbb{N}}$ is a sequence of independent standard normal variables, we deduce from Theorem 1.2.26 and Theorem 1.2.27 that the random series

$$\sum_{n\geq 1} X_n \frac{2\sqrt{2}\sin\big((2n-1)\pi t/2\big)}{(2n-1)\pi}$$

converge a.s. in $L^{1,2}$ and, in particular uniformly on $[0,1]$. The limit is the Brownian motion. □

We conclude this subsection with a simple application of Mercer kernels that we will use in the future.

Proposition 1.2.29. *Let (M,g) be a smooth compact connected m-dimensional manifold and*

$$K : M\times M \to \mathbb{R}$$

a Mercer kernel on M. Set $\mu := \mathrm{vol}_g$. Suppose that for some $\ell\in\mathbb{N}$ the operator $[K]_\mu$ induces a continuous operator

$$[K]_\mu : L^2(M,\mu)\to C^\ell(M).$$

(This happens if, e.g., $K\in C^\ell(M\times M)$.) Let $(\lambda_n)_{n\geq 1}$ be the nonzero eigenvalues of $[K]_\mu$ repeated according to their multiplicity and let (ψ_n) be an orthonormal system of eigenfunctions corresponding to these eigenvalues. Fix a sequence $(X_n)_{n\in\mathbb{N}}$ of independent standard normal random variables. Then the following hold.

(i) *For any $n\in\mathbb{N}$, $\psi_n\in C^\ell(M)$ and*

$$C := \sup_{n\in\mathbb{N}} \lambda_n \|\psi_n\|_{C^\ell(M)} < \infty. \tag{1.2.19}$$

(ii) *The random series*

$$\sum_{n\geq 1} X_n\lambda_n^2\psi_n \tag{1.2.20}$$

converges a.s. *in $C^\ell(M)$.*

Proof. (i) Note that

$$\psi_n = \frac{1}{\lambda_n}K\psi_n \in C^\ell(M).$$

If we denote by C the norm of the bounded operator $[K]_\mu : L^2(M,\mu)\to C^\ell(M)$ we deduce that

$$\lambda_n\|\psi_n\|_{C^\ell(M)} = \|[K]_\mu\psi_n\|_{C^\ell(M)} \leq C\|\psi_n\|_{L^2(M)} = C.$$

(ii) We deduce from (i) and Theorem 1.2.27(ii) that

$$\sum_{n\in\mathbb{N}} \lambda_n^2\|\psi_n\|_{C^\ell(M)} \leq C\sum_{n\in\mathbb{N}}\lambda_n < \infty.$$

The conclusion now follows from Proposition 1.1.57. □

Remark 1.2.30. Proposition 1.2.29 is more restrictive than Theorem 1.2.26, but it does not require the a priori knowledge that K is the covariance kernel of a Gaussian C^ℓ-function on M.

The covariance kernel of the Gaussian C^ℓ-function defined by (1.2.20) is K^{*4}, where K^{*n} is defined inductively as

$$K^{*n+1}(x,y) = \big(K^{*n} * K\big)(x,y) := \int_M K^{*n}(x,z)K(z,y)\mu\big[\,dy\,\big].$$

If we apply Theorem 1.2.26 to the kernel K^{*4} we obtain Proposition 1.2.29. However, we could do this only because Proposition 1.2.29 guarantees that K^{*4} is the covariance kernel of a Gaussian C^ℓ-function. □

Example 1.2.31 (Random Fourier series)**.** Consider the m-dimensional torus $\mathbb{T}^m := \big(\mathbb{R}/\mathbb{Z}\big)^m$ equipped with its flat metric. Denote by $\vec{\theta} = (\theta^1,\dots,\theta^m) \in (\mathbb{R}/\mathbb{Z})^m$ the resulting angular coordinates. The Laplacian[4] of the flat metric $g_1 = (d\theta^1)^2 + \cdots + (d\theta^m)^2$ on $\mathbb{T}^m$ has the form

$$\Delta = -\sum_{i=1}^m \partial^2_{\theta^i}.$$

We set $u_0 = 1$ and, for $\vec{\ell} = (\ell_1,\dots,\ell_m) \in \mathbb{Z}^m \setminus \{0\}$, we define $u_{\vec{\ell}}, v_\ell : \mathbb{R}^m \to \mathbb{R}$

$$u_{\vec{\ell}}(\vec{\theta}) = \sqrt{2}\cos 2\pi\langle\vec{\ell},\vec{\theta}\rangle,\;\; v_{\vec{\ell}} = \sqrt{2}\sin 2\pi\langle\vec{\ell},\vec{\theta}\rangle.$$

These functions are eigenfunctions of the Laplacian operator

$$\Delta = -\sum_{j=1}^n \partial^2_{\theta^j}.$$

More precisely,

$$\Delta u_{\vec{\ell}} = |2\pi\vec{\ell}|^2 u_{\vec{\ell}},\;\; \Delta v_{\vec{\ell}} = |2\pi\vec{\ell}|^2 v_{\vec{\ell}},\;\; |\vec{\ell}|^2 = \sum_{i=1}^m \ell_i^2.$$

Consider as in Example 1.2.17 the lexicographic order $\prec$ on $\mathbb{R}^m$. The collection

$$\big\{\, u_{\vec{\ell}},\; v_{\vec{k}};\;\; \vec{\ell},\vec{k} \in \mathbb{Z}^m,\; \vec{k} \succ 0,\;\; \vec{\ell} \succeq 0\,\big\}$$

is complete orthonormal system of $L^2(\mathbb{T}^m)$.

Pick an even *Schwartz function* $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$ such that $\mathfrak{a}(0) = 1$. We will refer to such a function as *amplitude*. For $R > 0$ (meant to be large) set

$$F^R(\vec{\theta}) = F^R_{\mathfrak{a}}(\vec{\theta}) = R^{-m/2}\Big(\,A_0u_0 + \sum_{\vec{\ell}\succ 0} \mathfrak{a}\big(\,|\,2\pi\vec{\ell}\,|/R\big)\big(\,A_{\vec{\ell}}u_{\vec{\ell}}(\vec{\theta}) + B_{\vec{\ell}}v_{\vec{\ell}}(\vec{\theta})\,\big)\,\Big), \tag{1.2.21}$$

where $A_{\vec{\ell}}$, $B_{\vec{k}}$ are independent standard normal random variables. Since $\mathfrak{a}$ is even, the function $\boldsymbol{b}(t) := \mathfrak{a}\big(\sqrt{|t|}\,\big)$ is also Schwartz so $\mathbb{R}^m \ni \xi \mapsto \mathfrak{a}\big(\,|\xi|\,\big) = \boldsymbol{b}\big(\,|\xi|^2\,\big) \in \mathbb{R}$ is Schwartz and $O(m)$-invariant.

[4]Throughout this book the Laplacian is the geometers' Laplacian and it is a nonnegative operator.

Since $\mathfrak{a}$ is a Schwartz function we deduce from Proposition 1.1.57 that the above series converges a.s. in any $C^k(\mathbb{T})$. If we define $Z_{\vec{\ell}}$ by

$$Z_{\vec{\ell}} := \begin{cases} A_0, & \vec{\ell} = 0,\\ \frac{1}{\sqrt{2}}\big(A_{\vec{\ell}} - \boldsymbol{i}B_{\vec{\ell}}\big), & \vec{\ell} \succ 0,\\ \bar{Z}_{-\vec{\ell}}, & \vec{\ell} \prec 0.\end{cases} \tag{1.2.22}$$

then we have

$$F^R_{\mathfrak{a}}(\vec{\theta}) = R^{-m/2}\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,\big|2\pi\vec{\ell}\big|/R\,\big)Z_{\vec{\ell}}e^{2\pi\boldsymbol{i}\langle\vec{\ell},\theta\rangle}. \tag{1.2.23}$$

Since $\mathfrak{a}\big(\,|2\pi\vec{\ell}|/R\,\big)$ decays very fast as $|\vec{\ell}| \to \infty$ we deduce from Kolmogorov's two-series theorem that for any $\nu \in \mathbb{N}$ the random series

$$\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,|2\pi\vec{\ell}|/R\,\big)^2\,\|\,e_{\vec{\ell}}\|^2_{C^\nu(\mathbb{T}^m)}$$

converges a.s. and thus the series

$$\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,|2\pi\vec{\ell}|/R\,\big)Z_{\vec{\ell}}e_{\vec{\ell}}$$

converges a.s. in $C^\nu\big(\,\mathbb{T}^m\,\big)$. In particular, this shows that the Gaussian function $F^R_{\mathfrak{a}}$ is a.s. smooth. Its covariance kernel is

$$\mathcal{C}^R_{\mathfrak{a}}\big(\,\vec{\varphi}+\vec{\tau},\vec{\varphi}\,\big) = C^R_{\mathfrak{a}}(\vec{\tau}) = R^{-m}\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,|2\pi\vec{\ell}|/R\,\big)^2 e^{2\pi\boldsymbol{i}\langle\vec{\ell},\vec{\tau}\rangle}. \tag{1.2.24}$$

Define $w_{\mathfrak{a}} = w_{\mathfrak{a},m} : \mathbb{R}^m \to \mathbb{R}$, $w_{\mathfrak{a}}(\xi) = \mathfrak{a}\big(\,|\xi|\,\big)^2$, and denote by $\widehat{w}_{\mathfrak{a}}$ the Fourier transform of $w_{\mathfrak{a}}$,

$$\widehat{w}_{\mathfrak{a}}(\boldsymbol{x}) = \int_{\mathbb{R}^m} e^{-\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mathfrak{a}\big(\,|\xi\,\big)^2 d\xi.$$

Using Poisson's summation formula (B.2.6) we deduce

$$\mathcal{C}^R_{\mathfrak{a}}(\vec{\varphi}+\vec{\tau},\vec{\varphi}) = \frac{1}{(2\pi)^m}\sum_{\vec{k}\in\mathbb{Z}^m}\widehat{w}_{\mathfrak{a}}\big(\,(\vec{k}-\vec{\tau})R\,\big). \tag{1.2.25}$$

Observe that $\mathcal{C}^R_{\mathfrak{a}}$ is the Schwarz kernel of the smoothing operator $\mathfrak{a}\big(\,\hbar\sqrt{\Delta}\,\big)^2$, $\hbar = R^{-1}$, and thus the associated Gaussian function is a.s. smooth.

For example if $\mathfrak{a}(t) = e^{-t^2/4}$, then $w_{\mathfrak{a}} = e^{-|\xi|^2/2}$ and we deduce from Proposition 1.1.15 that

$$\widehat{w}_{\mathfrak{a}}(\boldsymbol{x}) = (2\pi)^{m/2}e^{-|\xi|^2/2},\;\; \mathcal{C}^R_{\mathfrak{a}}(\vec{\theta},\vec{\varphi}) = \sum_{\vec{k}\in\mathbb{Z}^m} e^{-R^2|\vec{k}-\vec{\tau}|^2/2}$$

We can think of $F^R_{\mathfrak{a}}$ either as a function on $\mathbb{T}^m$, or as a $\mathbb{Z}^m$-periodic function of $\mathbb{R}^m$. If we formally let $R \to \infty$ in the equality

$$R^{m/2}F^R_{\mathfrak{a}}(\vec{\theta}) = \sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,|2\pi\vec{\ell}|/R\,\big)Z_{\vec{\ell}}e_{\vec{\ell}}(\vec{\theta})$$

we deduce

$$W_\infty(\vec{\theta}) \stackrel{?}{=} \lim_{R\to\infty} R^{m/2}F^R_{\mathfrak{a}}(\vec{\theta}) = \sum_{\vec{\ell}\in\mathbb{Z}^m} Z_{\vec{\ell}}e_{\vec{\ell}}(\vec{\theta}).$$

The series on the right-hand-side is a.s. divergent but we can still assign a meaning to W_∞ as a random generalized function, i.e., a random linear functional

$$C^\infty(\mathbb{T}^m) \to \mathbb{R},\ \ W_\infty(f) = \sum_{\vec{\ell}\in\mathbb{Z}^m} Z_{\vec{\ell}}\big(\, f, e_{\vec{\ell}}(\vec{\theta}\,)\,\big)_{L^2(\mathbb{T}^m)}.$$

A simple computation shows that for any functions $f_0, f_1 \in C^\infty(\mathbb{T}^m)$

$$\operatorname{Cov}\big[\, W_\infty(f_0), W_\infty(f_1)\,\big] = \sum_{\vec{\ell}\in\mathbb{Z}^m} \big(\, f_0, e_{\vec{\ell}}\,\big)_{L^2(\mathbb{T}^m, g_1)}\big(\, f_1, e_{\vec{\ell}}\,\big)_{L^2(\mathbb{T}^m, g_1)} = \big(\, f_0, f_1\,\big)_{L^2(\mathbb{T}^m, g_1)}.$$

The last equality shows that W_∞ is the Gaussian white noise on $\mathbb{T}^m$ driven by the volume measure vol_{g_1}; see [**67**]. In other words, one could think of the family $\big(\, W_R = R^{m/2}F^R_{\mathfrak{a}}\,\big)_{R>0}$ as a white noise approximation.

Here is another more geometric way of constructing $F^R_{\mathfrak{a}}$. For $R > 0$ meant to be large, we denote by Δ_R the Laplacian of the metric $g_R = R^2 g_1$. Observe that

$$\operatorname{vol}\big[\, M, g_R\,\big] = R^m \operatorname{vol}\big[\, M, g_1\,\big] = R^m,\ \ \Delta_R = R^{-2}\Delta_1.$$

Note that the torus $(\mathbb{T}^m, g_R)$ is *isometric* to the torus $\mathbb{R}^m/(R\mathbb{Z})^m$ so as $R \to \infty$ it starts to resemble[5] more and more like $\mathbb{R}^n$ with the canonical metric. Set

$$u^R_{\vec{k}} = R^{-m/2}u_{\vec{k}},\ \ v^R_{\vec{\ell}} = R^{-m/2}v_{\vec{\ell}}.$$

The collection

$$\big\{\, u^R_{\vec{k}},\ v^R_{\vec{\ell}};\ \vec{k} \succeq 0,\ \ \vec{\ell} \succ 0\,\big\}$$

is a complete $L^2(M, g_R)$-orthonormal system of real eigenfunctions of Δ_R. Moreover

$$\Delta_R u^R_{\vec{k}} = \lambda_{\vec{k}}(R),\ \ \Delta_R v^R_{\vec{\ell}} = \lambda_{\vec{\ell}}(R) v^R_{\vec{\ell}},\ \ \lambda_{\vec{k}}(R) = R^{-2}\big|\, 2\pi\vec{k}\,\big|^2$$

Then

$$F^R_{\mathfrak{a}}(\vec{\theta}) = \mathfrak{a}(0)A_0 u^R_0\big(\,\vec{\theta}\,\big) + \sum_{\vec{\ell}\succ 0} \mathfrak{a}\big(\, \lambda_{\vec{\ell}}(R)^{1/2}\,\big)\big(\, A_{\vec{\ell}} u^R_{\vec{\ell}}\big(\,\vec{\theta}\,\big) + B_{\vec{\ell}} v^R_{\vec{\ell}}\big(\,\vec{\theta}\,\big)\,\big).$$

Let me give an idea of the statistical meaning of the large parameter R.

Suppose for example that $\mathfrak{a}$ is supported on the interval $[-1, 1]$ and even better, it is a smooth approximation of the (discontinuous) indicator function $\boldsymbol{I}_{[-1,1]}$. Then the random series (1.2.21) is a random finite linear combination of eigenfunctions of the Laplacian on $\mathbb{T}^m$ corresponding to the eigenvalues satisfying $\sqrt{\lambda} \leq R$. To put it differently, the random function $R^{m/2}F^R_{\mathfrak{a}}$ defines a Gaussian measure Γ_R on the Féchet space $C^\infty(\mathbb{T}^m)$ and the vector space spanned by the eigenfunctions corresponding to the eigenvaluess $\lambda \leq R^2$ is contained in the support of Γ_R. As $R \to \infty$ the the support of Γ_R increases and it covers more and more of the space $C^\infty(\mathbb{T}^m)$. Moreover, since $\mathfrak{a}(\lambda/R) \to \mathfrak{a}(0)$ as $R \to \infty$ some of the bias towards eigenfunctions corresponding smaller eigenvalues built-in the the definition of $F^R_{\mathfrak{a}}$ starts to dissipate and, intuitively, in the white noise limit we reach an unbiased sampling of all the smooth functions on $\mathbb{T}^m$. This last claim is only a nonrigorous guiding motivation.

Finally let me give a third, functional analytic description of the function $F^R_{\mathfrak{a}}$.

As $R \to \infty$ the Gaussian measure Γ_R converges in some sense to Γ_∞, the *Gaussian white noise*. This white noise in fact a measure of $C^{-\infty}(\mathbb{T}^m)$, the topological dual of the Fréchet space $C^\infty(\mathbb{T}^m)$. The elements of $C^{-\infty}(\mathbb{T}^m)$ are commonly known as *generalized functions,* or

[5] For centuries people thought that Earth was flat, i.e., it resembled $\mathbb{R}^2$.

distributions. For a more in depth look at this aspect we refer to [67, Chap. III]. Consider the smoothing operator $\mathfrak{a}\big(\hbar\sqrt{\Delta}\big)$, $\hbar = R^{-1}$. Then

$$R^{m/2}F^R_{\mathfrak{a}} = \mathfrak{a}\big(\hbar\sqrt{\Delta}\big)W$$

where W is a generalized function with distribution Γ_∞. Note that when $\mathfrak{a}(x) = e^{-x^2}$ and $t = \hbar^{1/2}$, then $\mathfrak{a}\big(\hbar\sqrt{\Delta}\big) = e^{-t\Delta}$ - the heat operator. □

1.2.5. Stationary and isotropic Gaussian fields. Fix a centered, complex valued, random function $F : \mathbb{R}^m \to \mathbb{C}$ that is L^2-continuous, i.e.,

$$\lim_{\boldsymbol{s}\to\boldsymbol{t}} \big\| F(\boldsymbol{s}) - F(\boldsymbol{t}) \big\|_{L^2} = 0, \;\; \forall \boldsymbol{t} \in \mathbb{R}^m.$$

In particular, the covariance kernel of F,

$$\mathcal{K} = \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{C}, \;\; \mathcal{K}(\boldsymbol{x},\boldsymbol{y}) = \mathbb{E}\big[\, F(\boldsymbol{x})\bar{F}(\boldsymbol{y})\,\big]$$

is continuous.

Suppose that G is a Lie group that acts on $\mathbb{R}^m$,

$$G \times \mathbb{R}^m \to \mathbb{R}^m, \;\; G \times \mathbb{R}^m \ni (g,\boldsymbol{x}) \to g\cdot\boldsymbol{x} \in \mathbb{R}^m.$$

We say that F is *G-invariant* if for any $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n \in \mathbb{R}^m$ and any $g \in G$ the random vectors

$$\big(\, F(g\cdot\boldsymbol{x}_1),\dots,F(g\cdot\boldsymbol{x}_n)\,\big) \;\text{ and }\; \big(\, F(\boldsymbol{x}_1),\cdots,F(\boldsymbol{x}_n)\,\big)$$

have identical distributions.

A necessary condition for this to happen is

$$\mathcal{K}\big(\, g\cdot\boldsymbol{x}, g\cdot\boldsymbol{y}\,\big) = \mathcal{K}\big(\,\boldsymbol{x},\boldsymbol{y}\,\big), \;\; \forall g \in G, \;\; \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^m.$$

The first interesting case is when G is the group of translations $G \cong \big(\mathbb{R}^m, +\big)$. The centered random function F is called *homogeneous* or *stationary* if it is invariant with respect to the group of translations, i.e., for any $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n \in \mathbb{R}^m$ and any $\boldsymbol{t} \in \mathbb{R}^m$ the random vectors

$$\big(\, F(\boldsymbol{t}+\boldsymbol{x}_1),\dots,F(\boldsymbol{t}+\boldsymbol{x}_n)\,\big) \;\text{ and }\; \big(\, F(\boldsymbol{x}_1),\cdots,F(\boldsymbol{x}_n)\,\big)$$

have identical distributions. In particular

$$\mathcal{K}\big(\,\boldsymbol{t}+\boldsymbol{x},\boldsymbol{t}+\boldsymbol{y}\,\big) = \mathcal{K}\big(\,\boldsymbol{x},\boldsymbol{y}\,\big), \;\; \forall \boldsymbol{t},\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^m.$$

This happens iff and only if there exists a continuous function $K : \mathbb{R}^m \to \mathbb{C}$ such that

$$\mathcal{K}(\boldsymbol{x},\boldsymbol{y}) = K(\boldsymbol{x}-\boldsymbol{y}), \;\; \forall \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^m. \tag{1.2.26}$$

A centered random function on $\mathbb{R}^m$ is called *wide sense stationary* if its covariance kernel satisfies (1.2.26). This imposes severe restrictions on K because for any $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n \in \mathbb{R}^n$ the hermitian $n\times n$-matrix $\big(\, K(\boldsymbol{x}_i - \boldsymbol{x}_j)\,\big)_{1\le i,j\le n}$ has to be nonnegative definite.

The continuous functions $K : \mathbb{R}^m \to \mathbb{C}$ with this property are called *nonnegative definite.* They have a Fourier theoretic characterization.

Theorem 1.2.32 (S. Bôchner)**.** *Let $K : \mathbb{R}^m \to \mathbb{C}$ be a continuous function. The following are equivalent.*

(i) *The function K is nonnegative definite.*

(ii) *There exists a finite Borel measure* μ *on* $\mathbb{R}^m$ *such that*

$$K(\boldsymbol{x}) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \xi, \boldsymbol{x}\rangle}\mu\big[\,d\xi\,\big]$$

□

For a proof we refer to [**136**, I.24], [**138**, Sec. 1.4] or [**146**, Thm. 9.17]. The measure μ above is uniquely determined by the function K via the inverse Fourier transform in the space of tempered distributions. It is called the *spectral measure* of the wide sense stationary random function.

Observe that if F is a centered L^2-continuous real *Gaussian* function, then F is stationary iff it is wide sense stationary.

Example 1.2.33. Suppose that Z is a centered symmetric complex[6] Gaussian random variable and $\Phi : \mathbb{R}^m \to \mathbb{C}$ is a nonzero continuous function. We obtain a random function $F(\boldsymbol{x}) = Z\Phi(\boldsymbol{x})$. A simple computation shows that if this function is wide sense stationary iff there exist $\xi \in \mathbb{R}^m$ and $A \in \mathbb{C} \setminus \{0\}$ such that $\Phi(\boldsymbol{x}) = Ae^{\boldsymbol{i}\langle \xi, \boldsymbol{x}\rangle}$; see [**161**, Sec.7] for details. The covariance kernel of this function is

$$\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}) = \big|\,A\,\big|^2 e^{\boldsymbol{i}\langle \xi, \boldsymbol{x}-\boldsymbol{y}\rangle}.$$

The spectral measure is $\big|\,A\,\big|^2\delta_\xi$.

Consider now a simple linear combination of random functions of the above type

$$G(\boldsymbol{x}) = Z_1 e^{\boldsymbol{i}\langle \xi_1, \boldsymbol{x}\rangle} + Z_2 e^{\boldsymbol{i}\langle \xi_2, \boldsymbol{x}\rangle}.$$

The random function G is wide sense stationary iff $\mathbb{E}\big[\,Z_1\bar{Z}_2\,\big] = 0$. In this case the spectral measure is

$$\mathbb{E}\big[\,|Z_1|^2\,\big]\delta_{\xi_1} + \mathbb{E}\big[\,|Z_2|^2\,\big]\delta_{\xi_2}.$$

The random function G is real valued iff $\xi_2 = -\xi_1$, $Z_2 = \bar{Z}_1$. In this case

$$G(t) = X_1\cos\langle \xi_1, \boldsymbol{x}\rangle + Y_1\sin\langle \xi_1, \boldsymbol{x}\rangle,\;\; Z_1 = \frac{1}{2}\big(\,X_1 - \boldsymbol{i}Y_1\,\big).$$

□

Example 1.2.34. Consider the Gaussian real function $F^R_{\mathfrak{a}}$ defined by (1.2.21) discussed in Example 1.2.31. We recall that $\mathfrak{a} : \mathbb{R} \to \mathbb{R}$ is an amplitude, i.e., an even Schwartz function such that $\mathfrak{a}(0) = 1$ and

$$\begin{aligned} F^R_{\mathfrak{a}}(\vec{\theta}) &= R^{-m/2}\Big(\,A_0u_0 + \sum_{\vec{\ell}\succ 0}\mathfrak{a}\big(\,\big|\,2\pi\hbar\vec{\ell}\,\big|\big)\big(\,A_{\vec{\ell}}u_{\vec{\ell}}(\vec{\theta}) + B_{\vec{\ell}}v_{\vec{\ell}}(\vec{\theta})\,\big)\,\Big) \\ &= R^{-m/2}\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,\big|\,2\pi\vec{\ell}\,\big|/R\,\big)Z_{\vec{\ell}}e_{\vec{\ell}}(\vec{\theta}),\;\; e_{\vec{\ell}}(\vec{\theta}) := e^{2\pi\boldsymbol{i}\langle\vec{\ell},\vec{\theta}\rangle}. \end{aligned}$$

We think of $F^R_{\mathfrak{a}}$ as a function on $\mathbb{R}^m$ that is periodic with respect to the lattice $\mathbb{Z}^m$. Equivalently, we can think of it as a function on the m-dimensional torus $\mathbb{T}^m = \mathbb{R}^m/\mathbb{Z}^m$. We have seen in in Example 1.2.31 that is a.s. smooth. Its covariance kernel is given by (1.2.24)

$$\mathcal{C}^R_{\mathfrak{a}}(\vec{\theta}, \vec{\varphi}) = R^{-m}\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,\big|\,2\pi\vec{\ell}\,\big|/R\,\big)^2 e_{\vec{\ell}}(\vec{\theta} - \vec{\varphi}).$$

[6]See Definition 1.1.37.

Hence $F_{\mathfrak{a}}^R$ is stationary and $C_{\mathfrak{a}}^R(\vec{\tau}) = \mathcal{C}_{\mathfrak{a}}^R(\vec{\tau}+\vec{\varphi},\vec{\varphi})$ is given by

$$C_{\mathfrak{a}}^{\hbar}(\vec{\tau}) = R^{-m}\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,\big|\,2\pi\vec{\ell}\,\big|/R\,\big)^2 e_{\vec{\ell}}(\vec{\tau}).$$

If we set

$$\mu\big[\,d\xi\,\big] = \mu_{\mathfrak{a},R}\big[\,d\xi\,\big] := R^{-m}\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,\big|\,2\pi\vec{\ell}\,\big|/R\,\big)^2\delta_{2\pi\vec{\ell}}\big[\,d\xi\,\big],$$

we deduce that

$$K_{\mathfrak{a}}^R(\vec{\tau}) = \int_{\mathbb{R}^n} e^{-\boldsymbol{i}\langle\xi,\vec{\tau}\rangle}\mu_{\mathfrak{a},R}\big[\,d\xi\,\big] = \int_{\mathbb{R}^n} e^{\boldsymbol{i}\langle\xi,\vec{\tau}\rangle}\mu_{\mathfrak{a},R}\big[\,d\xi\,\big].$$

Thus $\mu_{\mathfrak{a},R}$ is the spectral measure of this homogeneous random function.

We deduce from (1.2.25) that

$$C_{\mathfrak{a}}^R(\vec{\tau}) = \sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,\big|\,2\pi\vec{\ell}\,\big|/R\,\big)^2 e^{-2\pi\boldsymbol{i}\langle\vec{\ell},\vec{\tau}\rangle} = \sum_{\vec{k}\in\mathbb{Z}^m}\boldsymbol{K}_{\mathfrak{a}}\big(\,(\vec{k}-\vec{\tau})R\,\big),$$

where

$$\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}) = \frac{1}{(2\pi)^m}\widehat{w_{\mathfrak{a}}}(\boldsymbol{x}) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{-\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mathfrak{a}\big(\,|\xi|\,\big)^2 d\xi.$$

We can rewrite this in a more conceptual form.

We introduce the lattice $\Lambda_R = \big(\,2\pi R^{-1}\mathbb{Z}\,\big)^m$ and its dual $L_R = \big(\,R\mathbb{Z}\,\big)^m$. We set $\boldsymbol{x} := \vec{\tau}R$ and we deduce

$$C_{\mathfrak{a}}^R(R^{-1}\boldsymbol{x}) = R^{-m}\sum_{\omega\in\Lambda_R}\mathfrak{a}\big(\,\omega\,\big)^2 e^{\boldsymbol{i}\langle\omega,\boldsymbol{x}\rangle} = \sum_{\boldsymbol{t}\in L_R}\boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{t}-\boldsymbol{x}\,\big). \tag{1.2.27}$$

Note that $\mathcal{K}_{\mathfrak{a}}^R\big(\,\boldsymbol{x},\boldsymbol{y}\,\big) := C_{\mathfrak{a}}^R\big(\,R^{-1}(\boldsymbol{x}-\boldsymbol{y})\,\big)$ is the covariance kernel of the stationary Gaussian function

$$\Psi_{\mathfrak{a}}^R(\boldsymbol{x}) = F_{\mathfrak{a}}^R\big(\,R^{-1}\boldsymbol{x}\,\big)$$

that is periodic with respect to the lattice $L_R = (R\mathbb{Z})^m$. Set

$$\boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{x}) := \mathcal{K}_{\mathfrak{a}}^R\big(\,0,\boldsymbol{x}\,\big) = C_{\mathfrak{a}}^R\big(\,R^{-1}\boldsymbol{x}\,\big).$$

We have

$$\boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{x}) - \boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}) = \sum_{\boldsymbol{t}\in L_R\setminus\{0\}}\boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{x}-\boldsymbol{t}\,\big).$$

Since $\widehat{w_{\mathfrak{a}}}$ is a Schwartz function we deduce that

$$\lim_{R\to\infty}\boldsymbol{K}_{\mathfrak{a}}^R = \boldsymbol{K}_{\mathfrak{a}} \quad \text{in } C^k\big(\,\mathbb{R}^m\,\big),\ \forall k\in\mathbb{N}. \tag{1.2.28}$$

More precisely, for every ball $B\subset\mathbb{R}^m$, every $k\in\mathbb{N}$, and every $N>0$ there exists $C=C(k,N,B)>0$ such that

$$\forall R>1:\ \ \big\|\,\boldsymbol{K}_{\mathfrak{a}}^R - \boldsymbol{K}_{\mathfrak{a}}\,\big\|_{C^k(B)} \le CR^{-N}. \tag{1.2.29}$$

□

The Gaussian function $F:\mathbb{R}^m\to\mathbb{R}$ is called *isotropic* if it is homogeneous and invariant with respect to the natural action of the orthogonal group $O(m)$ on $\mathbb{R}^m$. If $\mathcal{K}_F$ is the covariance kernel of F then there exists a one-variable function K_F such that

$$\mathcal{K}_F\big(\,\boldsymbol{x},\boldsymbol{y}\,\big) = K_F\big(\,|\boldsymbol{x}-\boldsymbol{y}|\,\big),\ \ \forall\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^m.$$

Example 1.2.35. Suppose that $\mathfrak{a}:\mathbb{R}\to\mathbb{R}$ is an even Schwartz function such that $\mathfrak{a}(0)=1$.

Consider the finite Borel measure $\mu\in\mathrm{Meas}(\mathbb{R}^m)$

$$\mu\big[\,d\xi\,\big]=\mu_{\mathfrak{a}}\big[\,d\xi\,\big]=\frac{1}{(2\pi)^m}w_{\mathfrak{a},m}\big(\,\xi\,\big)\boldsymbol{\lambda}\big[\,d\xi\,\big],\;\; w_{\mathfrak{a},m}\big(\,\xi\,\big)=\mathfrak{a}\big(\,|\xi|\,\big)^2.$$

Its characteristic function is the nonnegative definite function

$$\boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{x}\,\big)=\int_{\mathbb{R}^m}e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mu\big[\,d\xi\,\big]=\frac{1}{(2\pi)^m}\widehat{w_{\mathfrak{a}}}(\boldsymbol{x})=\frac{1}{(2\pi)^m}\int_{\mathbb{R}^m}e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mathfrak{a}\big(\,|\xi|\,\big)^2\boldsymbol{\lambda}\big[\,d\xi\,\big].\tag{1.2.30}$$

Clearly $\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x})$ is an $O(m)$-invariant, real valued Schwartz function. Then $\boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{x}-\boldsymbol{y}\,\big)$ is the covariance kernel of a real valued, smooth isotropic Gaussian function $\Phi=\Phi_{\mathfrak{a}}$ on $\mathbb{R}^m$ with spectral measure $\mu_{\mathfrak{a}}$.

A good example to have in mind is $\mathfrak{a}(t)=e^{-t^2/4}$. Then $\mathfrak{a}(t)^2=e^{-t^2/2}$, and

$$\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x})=\frac{1}{(2\pi)^m}\int_{\mathbb{R}^m}e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}e^{-\frac{|\xi|^2}{2}}d\xi=\frac{1}{(2\pi)^{m/2}}e^{-\frac{|\boldsymbol{x}|^2}{2}}.$$

Thus $\boldsymbol{K}_{\mathfrak{a}}$ is in this case the density of the canonical Gaussian measure $\boldsymbol{\Gamma}_{\mathbb{1}}$ on $\mathbb{R}^m$. In this case $\Psi(t)=\frac{1}{(2\pi)^{m/2}}e^{-t/2}$.

Since $w_{\mathfrak{a},m}>0$ in an open neighborhood of the origin, we deduce from [**157**, Thm. 6.8] that if $\boldsymbol{x}_1,\dots,\boldsymbol{x}_N\in\mathbb{R}^m$ are distinct points, then the symmetric $N\times N$ matrix

$$\big(\,\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}_i-\boldsymbol{x}_j)\,\big)_{1\leq i,j\leq N}$$

is *positive* definite. This matrix is the variance matrix of the Gaussian vector

$$\big(\,\Phi_{\mathfrak{a}}(\boldsymbol{x}_1),\dots,\Phi_{\mathfrak{a}}(\boldsymbol{x}_N)\,\big).$$

Hence, for any distinct points $\boldsymbol{x}_1,\dots,\boldsymbol{x}_N\in\mathbb{R}^n$, the above Gaussian vector is nondegenerate. In other words, $\Phi_{\mathfrak{a}}$ is ∞-ample in the sense of Definition 1.2.8.

Observe that for any multi-indices $\alpha\in\big(\,\mathbb{Z}_{\geq 0}\,\big)^m$, $|\alpha|=|\beta|$, we have

$$\mathbb{E}\big(\,\partial^\alpha\Phi_{\mathfrak{a}}(\boldsymbol{x})\partial^\beta\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big)=\partial^\alpha_x\partial^\beta_y\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}-\boldsymbol{y}\,)\big|_{\boldsymbol{x}=\boldsymbol{y}}$$

$$=\int_{\mathbb{R}^m}\boldsymbol{\xi}^\alpha\boldsymbol{\xi}^\beta\mu_{\mathfrak{a}}\big[\,d\boldsymbol{\xi}\,\big],\;\;\boldsymbol{\xi}^\alpha:=\xi_1^{\alpha_1}\cdots\xi_m^{\alpha_m}$$

This shows that for any $k\in\mathbb{N}$ and any $\boldsymbol{x}\in\mathbb{R}^n$ the variance the Gaussian vector $\big(\,\partial^\alpha\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big)_{|\alpha|=k}$ is the Gramian matrix of the functions $\big(\,\xi^\alpha\,\big)_{|\alpha|=k}$ with respect to the inner product in $L^2\big(\,\mathbb{R}^m,\mu_{\mathfrak{a}}\,\big)$. Since $\mathfrak{a}(0)=1$ we deduce that the functions ξ^α are linearly independent in $L^2\big(\,\mathbb{R}^m,\mu_{\mathfrak{a}}\,\big)$ so the determinant of their Gramian matrix is nonzero. Hence the Gaussian vector

$$\Phi_{\mathfrak{a}}(\boldsymbol{x})\oplus D\Phi_{\mathfrak{a}}(\boldsymbol{x})\oplus\cdots\oplus D^k\Phi_{\mathfrak{a}}(\boldsymbol{x})$$

is nondegenerate, for any $k\in\mathbb{N}$ and any $\boldsymbol{x}\in\mathbb{R}^m$. Above, $D^j\Phi_{\mathfrak{a}}(\boldsymbol{x})$ denotes the j-th order differential of $\Phi_{\mathfrak{a}}$ at $\boldsymbol{x}\in\mathbb{R}^m$. In other words $\Phi_{\mathfrak{a}}$ is J_k-ample for any $k\in\mathbb{N}$.

In Example 1.2.34 above we proved that the $\big(\,R\mathbb{Z}\,\big)^m$-periodic function

$$\Psi^R_{\mathfrak{a}}(\boldsymbol{x})=F^R_{\mathfrak{a}}\big(\,R^{-1}\boldsymbol{x}\,\big)$$

converges in distribution to the smooth isotropic function $\Phi_{\mathfrak{a}}$ as $R\to\infty$, i.e., the covariance kernel of $\Psi^R_{\mathfrak{a}}$ converges in C^∞ to the covariance kernel of $\Phi_{\mathfrak{a}}$ as $R\to\infty$.

For $R > 0$ we set

$$\mathfrak{a}_R(t) := \mathfrak{a}\big(t/R\big), \;\; \forall t \in \mathbb{R}.$$

Consider the finite Borel measure $\mu \in \mathrm{Meas}(\mathbb{R}^m)$

$$\mu_{\mathfrak{a}}^R\big[\,d\boldsymbol{\xi}\,\big] = \frac{1}{(2\pi)^m} w_{\boldsymbol{a},m}\big(\,R^{-1}\boldsymbol{\xi}\,\big)\boldsymbol{\lambda}\big[\,d\boldsymbol{\xi}\,\big] = \frac{1}{(2\pi)^m}\mathfrak{a}\big(\,|\boldsymbol{\xi}|/R\,\big)^2\boldsymbol{\lambda}\big[\,d\boldsymbol{\xi}\,\big].$$

Its characteristic function is the nonnegative definite function

$$\boldsymbol{K}_{\mathfrak{a}}^R\big(\,\boldsymbol{x}\,\big) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle\boldsymbol{\xi},\boldsymbol{x}\rangle}\mathfrak{a}\big(\,|\boldsymbol{\xi}|/R\,\big)^2 d\xi. \tag{1.2.31}$$

We set $\omega := R^{-1}\boldsymbol{\xi}$ in (**??**) and we deduce

$$\boldsymbol{K}_{\mathfrak{a}}^R\big(\,\boldsymbol{x}\,\big) = \frac{R^m}{(2\pi)^m}\int_{\mathbb{R}^m} e^{\boldsymbol{i}R\langle\omega,\boldsymbol{x}\rangle}\mathfrak{a}\big(\,|\omega|\,\big)^2 d\omega,$$

so that

$$\boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{x}) = R^m\boldsymbol{K}_{\mathfrak{a}}\big(\,R\boldsymbol{x}\,\big).$$

We deduce that $\boldsymbol{K}_{\mathfrak{a}}^R\big(\,\boldsymbol{x}-\boldsymbol{y}\,\big)$ is the covariance kernel of the Gaussian function

$$\Phi_{\mathfrak{a}}^R(\boldsymbol{x}) := R^{m/2}\Phi_{\mathfrak{a}}\big(\,R\boldsymbol{x}\,\big).$$

We want to investigate the behavior of $\boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{x})$ as $R\to\infty$. For example, in the special case $\mathfrak{a}(t) = e^{-t^2/4}$ we have

$$\boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{x}) = \frac{1}{(2\pi\hbar^2)^{m/2}}e^{-\frac{|\boldsymbol{x}|^2}{2\hbar^2}}, \;\; \hbar = R^{-1}$$

This is the density of the Gaussian measure $\boldsymbol{\Gamma}_{\hbar^2\mathbb{1}}$ which converges to the Dirac measure δ_0 as $R\to\infty$.

Since $K_{\mathfrak{a}}(\boldsymbol{x})$ is $O(m)$-invariant and smooth it has the form $\Psi\big(\,|\boldsymbol{x}|^2\,\big)$ for some smooth function $\Psi : [0,\infty)\to\mathbb{R}$. According to Schoenberg's characterization theorem [**157**, Thm. 7.13], the function Ψ must be completely monotone. In particular, Ψ is non-increasing, nonnegative and convex, [**157**, Lemma.7.3]. Using the Fourier inversion formula we deduce

$$\int_{\mathbb{R}^m}\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x})d\boldsymbol{x} = \mathfrak{a}(0)^2 = 1.$$

This implies that $\boldsymbol{K}_{\mathfrak{a}}$ is the density of a probability measure on $\mathbb{R}^m$. The rescaled measures $\boldsymbol{K}_{\mathfrak{a}}^R\big(\,\boldsymbol{x}\,\big)dx$ converge weakly to the Dirac measure δ_0. To use a terminology favored by physicists, we have

$$\boldsymbol{K}_{\mathfrak{a}}^R\big(\,\boldsymbol{x}\,\big)\to\delta(\boldsymbol{x}),$$

where $\delta(\boldsymbol{x})$ is Dirac's mysterious Delta function. In particular,

$$\boldsymbol{K}_{\mathfrak{a}}^R\big(\,\boldsymbol{x}-\boldsymbol{y}\,\big)\to\delta(\boldsymbol{x}-\boldsymbol{y}).$$

In other words, as $R\to\infty$, the Gaussian random function $\boldsymbol{K}_{\mathfrak{a}}^R$ converges in some sense to a Gaussian random "function" $\Phi_{\mathfrak{a}}^\infty$ whose covariance kernel is $\boldsymbol{K}_{\mathfrak{a}}^\infty(\boldsymbol{x}-\boldsymbol{y}) = \delta(\boldsymbol{x}-\boldsymbol{y})$. This is the Gaussian noise on W on $\mathbb{R}^m$ driven by the Lebesgue measure.

Formally, W is a Gaussian random generalized function or, equivalently, a Gaussian probability measure on a space of generalized functions or distributions on $\mathbb{R}^m$. The white noise W is discussed in detail in [**67**, Sec.III.4].

We can provide an explicit description of $\Phi_{\mathfrak{a}}$ in terms of white noise W. Consider the geometers' Laplacian on $\mathbb{R}^m$,

$$\Delta = -\sum_{k=1}^{m} \partial_{x_k}^2.$$

For any Schwarz function $f \in \mathcal{S}(\mathbb{R}^m)$ we have

$$\Delta f(\boldsymbol{x}) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \boldsymbol{\xi}, \boldsymbol{x}\rangle} \widehat{f}(\boldsymbol{\xi}) d\boldsymbol{\xi}$$

where $\widehat{f}(\boldsymbol{\xi})$ denotes the Fourier transform of f. As mentioned earlier, we can write the amplitude $\mathfrak{a}$ as $\mathfrak{a}(t) = \boldsymbol{b}(t^2)$, $\boldsymbol{b} \in \mathcal{S}(\mathbb{R}^n)$.

We have an integral operator $\mathbf{Op}\left(\mathfrak{a}\right) := \mathfrak{a}\big(\sqrt{\Delta}\big) = \boldsymbol{b}\big(\Delta\big)$ defined by

$$\begin{aligned}\mathbf{Op}\left(\mathfrak{a}\right) f(x) = \mathfrak{a}\big(\sqrt{\Delta}\big) f(x) &= \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \boldsymbol{\xi}, \boldsymbol{x}\rangle} \boldsymbol{b}\big(|\boldsymbol{\xi}|^2\big) \widehat{f}\big(\boldsymbol{\xi}\big) d\boldsymbol{\xi} \\ &= \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \boldsymbol{\xi}, \boldsymbol{x}\rangle} \mathfrak{a}\big(|\boldsymbol{\xi}|\big) \widehat{f}\big(\boldsymbol{\xi}\big) d\boldsymbol{\xi}.\end{aligned}$$

Note that

$$\mathbf{Op}\left(\mathfrak{a}^2\right) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \boldsymbol{\xi}, \boldsymbol{x}\rangle} \mathfrak{a}\big(|\boldsymbol{\xi}|\big)^2 \widehat{f}\big(\boldsymbol{\xi}\big) d\boldsymbol{\xi}. = \int_{\mathbb{R}^m} \boldsymbol{K}_{\mathfrak{a}}\big(\boldsymbol{x} - \boldsymbol{y}\big) f\big(\boldsymbol{y}\big) d\boldsymbol{y}.$$

For example, if $\mathfrak{a}(t) = e^{-t^2/2}$, then $\mathbf{Op}(\mathfrak{a}) = e^{\frac{1}{2}\Delta}$. In general $\mathfrak{a}\big(\sqrt{\Delta}\big)$ is a smoothing operator, i.e., for any tempered distribution u we have

$$\mathfrak{a}\big(\sqrt{\Delta}\big) u \in C^{\infty}\big(\mathbb{R}^m\big).$$

If W is the Gaussian white noise mentioned above, then

$$\Phi_{\mathfrak{a}} = \mathfrak{a}\big(\sqrt{\Delta}\big) W.$$

For any $R > 0$ we set $\mathfrak{a}_R\big(t\big) = \mathfrak{a}(t/R)$. Note that $\mathfrak{a}_R$ converges in the sense of temperate distriibutions to the constant function 1. Its Fourier transform is the Dirac distribution in the $\boldsymbol{\xi}$-space. This translates into the fact that $\mathbf{Op}(\mathfrak{a}_R) \to \mathbb{1}$ as $R \to \infty$. Note that if $R = t^{-2}$, then

$$\mathbf{Op}\left(\mathfrak{a}_R\right) = e^{\frac{t}{2}\Delta}$$

and we recover the known fact that $e^{t\Delta} \to \mathbb{1}$ as $t \searrow 0$.

□

1.2.6. Gaussian random sections of a vector bundle. The concept of random section is not an artificial generalization. The main object of investigation of this book requires it. Suppose for example that M is a smooth, connected manifold and

$$\Phi : \Omega \times M \to \mathbb{R}$$

is a C^k random function on M, $k \geq 1$. Then the differential $d\Phi$ should be viewed as a random section of the cotangent bundle T^*M. Its zeros are the critical points of Φ

Consider a more general problem. Suppose that M is a smooth, compact, connected m-dimensional manifold and $\pi : E \to M$ is a smooth real vector bundle of rank r. For each $x \in M$, the fiber $E_x = \pi^{-1}(\{x\})$ of E has a natural structure of real vector space of dimension r.

From a set theoretic point of view, we can regard E as a family $\big(E_x\big)_{x\in M}$ of real vector spaces of dimension r. Loosely speaking, a random section of E is a family $\big(\Psi(x)\big)_{x\in M}$ of random vectors $\Psi(x) : (\Omega, \mathcal{S}, \mathbb{P}) \to E_x$.

This definition is not satisfactory since we are interested in regularity properties of random sections. We are interested only in Gaussian random sections so we take a different approach suggested by Example 1.2.18. This was pioneered by P. Baxendale [**15**]. For different but related approach we refer to [**114**, Sec. 1.2].

Denote by $C^k(E)$ the vector space of sections of E that are k-times continuously differentiable. We need to define on $C^k(E)$ a structure of separable Banach space and to do so we need to make some choices.

- Fix a smooth Riemannian metric g on M.
- Fix a smooth h metric on E. We denote by $(-,-)_{E_x}$ the induced inner product on E_x.
- Fix a connection (covariant derivative) ∇^h on E that is compatible with the metric h.

We will refer to a triplet (g, h, ∇) as above as a *geometric structure* on E. A *geometric vector bundle* is a vector bundle equipped with geometric structure. There are several geometric objects canonically induced by these choices; see [**117**, Sec. 3.3].

First, the metric g determines a a Borel measure vol_g on M, classically referred to as the *volume element* or the *volume density*. Next, the metric determines the *Levi-Civita connection* ∇^g on TM. The metric g also determines metrics on all the tensor bundles $TM^{\otimes p}\otimes(T^*M)^{\otimes q}$ and the connection ∇^g determines connections on these bundles compatible with the metrics induced by g. To ease the notational burden we will denote by ∇^g each of these connections.

Similarly, the metric h induces metrics in all the bundles $E^{\otimes p}\otimes(E^*)^{\otimes q}$ and the connection ∇^h determines connections on these bundles compatible with the induced metrics. We will denote by $\big|-\big|_x$ the Euclidean norms in any of the spaces $(T_x^*M)^{\otimes q}\otimes E^{\otimes p}$. We define the *jet bundle*[7]

$$J_k(E) := \bigoplus_{j-0}^{k} T^*M^{\otimes j}\otimes E. \tag{1.2.32}$$

The connections ∇^g and ∇^h induce a connection $\nabla = \nabla^{g,h}$ on the bundle $(T^*M)^{\otimes k}\otimes E$

$$\nabla : C^1\big((T^*M)^{\otimes k}\otimes E\big) \to C^0\big((T^*M)^{\otimes k+1}\otimes E\big).$$

We denote by ∇^q the composition

$$C^m(E) \stackrel{\nabla}{\to} C^{m-1}\big(T^*M\otimes E\big) \stackrel{\nabla}{\to} \cdots \stackrel{\nabla}{\to} C^1\big((T^*M)^{\otimes q-1}\otimes E\big) \stackrel{\nabla}{\to} C^0\big((T^*M)^{\otimes q}\otimes E\big).$$

For every section $\psi\in C^k(E)$ we define its k-th jet

$$J_k(\psi) = J_k(\psi,\nabla) = \bigoplus_{k=0}^{k}\nabla^k\psi$$

[7]The jet bundle can be defined invariantly without relying on choices of connections and, as such, its *is merely* an affine bundle. For the applications I have in mind I do not need such a generality. For details I refer to [**139**].

$$\|u\|_{C^k} = \sum_{j=0}^{q} \|\nabla^j \psi\|,$$

where

$$\|\nabla^j u\| = \sup_{x\in M} \left| \nabla^j u(x) \right|_x .$$

The norm $\| - \|_{C^k}$ depends on the standard choices, but different standard choices yield equivalent norms. The resulting normed space is a separable Banach space. Fix one such norm and denote by $C^k(E)$ the resulting separable Banach space.

For every $x \in M$ and $\boldsymbol{u}_x \in E_x$ we have evaluation maps

$$\mathbf{Ev}_x : C^k(E) \to E_x, \;\; \mathbf{Ev}_x(\psi) = \psi(x) \in E_x$$

and

$$\mathbf{Ev}_{x,\boldsymbol{u}_x} : C^k(E) \to \mathbb{R}, \;\; \mathbf{Ev}_{x,u_x}(\psi) = \big(\psi(x), u_x \big)_{E_x}.$$

The evaluation map $\mathbf{Ev}_{x,\boldsymbol{u}_x}$ is a continuous linear function and thus defines and element in the dual $C^k(E)^*$. Set

$$\mathcal{L} := \operatorname{span}\big\{ \mathbf{Ev}_{x,\boldsymbol{u}_x}; \;\; x \in M, \;\; \boldsymbol{u}_x \in E_x \big\}.$$

If we choose a dense countable set $\mathfrak{X} \subset M$ and for each $x \in \mathfrak{X}$ a basis $\{\boldsymbol{e}_1(x), \dots, \boldsymbol{e}_r(x)\}$ of E_x we deduce that the countable collection

$$\big\{ \mathbf{Ev}_{x,\boldsymbol{e}_i(x)}; \;\; x \in \mathfrak{X}, \;\; 1 \le i \le r \big\}$$

separates the points in $C^k(E)$ and, according to Blackwell's Theorem 1.1.41, it generates the Borel-sigma algebra of $C^k(E)$.

Definition 1.2.36. A *centered Gaussian measure* on $C^k(E)$ is a Borel probability measure Γ such that $\forall \xi \in \mathcal{L}$ the random variable $\xi : C^k(E) \to \mathbb{R}$ is centered Gaussian. □

Equivalently, if we denote by $\boldsymbol{T}$ the disjoint union

$$T = \bigcup_{x\in M} \{x\} \times E_x,$$

then Γ is centered Gaussian iff the random process

$$\mathcal{E}_\Gamma : \big(C^k(E), \Gamma \big) \times \boldsymbol{T} \to \mathbb{R}, \;\; (\psi; x, \boldsymbol{u}_x) \mapsto \mathbf{Ev}_{x,\boldsymbol{u}_x}(\psi)$$

is centered Gaussian. Corollary 1.1.46 shows that the measure Γ is uniquely determined by the distribution of the process Γ.

Inspired by statistical physicists, we will often refer to Gaussian measures on $C^k(E)$ as *Gaussian ensembles* of C^k sections.

Definition 1.2.37. Suppose that $E \to M$ is a smooth vector bundle over the smooth compact manifold of dimension m. Fix $n \ge 0$ and set $\boldsymbol{X} = C^n(E)$.

(i) A centered Gaussian C^n-section of E is a measurable map

$$\Psi : (\Omega, \mathcal{S}, \mathbb{P}) \to \big(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}} \big), \;\; \omega \mapsto \Psi_\omega$$

whose distribution $\Gamma = \mathbb{P}_\Psi$ is a Gaussian measure on $\boldsymbol{X}$.

(ii) Let $k \in \mathbb{N}$. The Gaussian section Ψ is called *k-ample* if for any distinct points $x_1, \dots x_k \in M$ the Gaussian vector

$$\Psi(p_1) \oplus \cdots \oplus \Psi(p_k) \in E_{x_1} \oplus \cdots \oplus E_{x_k}$$

is nondegenerate. The Gaussian section Ψ is said to be *ample* if it is 1-ample.

(iii) Let $k \leq n$. The Gaussian section Ψ is called *J_k-ample* if there exists a smooth connection ∇ on E such that, the associated k-th jet $J_k(\Psi, \nabla)$ is ample, i.e., for any $x \in M$, the Gaussian vector

$$J_k\big(\,\Psi(x)\,\big) \in J_k(E)_x = E_x \oplus T_x^*M \otimes E \oplus \cdots \oplus \big(\,T_x^*M^{\otimes k}\,\big) \otimes E \tag{1.2.33}$$

is nondegenerate.

□

Let us point out that if condition (iii) above holds for one smooth connection, then it holds for all smooth connections.

For each $x_0, x_1 \in M$ we have two Gaussian vectors $\Psi(x_i) : \Omega \to E_{x_i}$, $i = 0, 1$, and we define

$$\mathcal{K}(x_1, x_0) := \operatorname{Cov}\big[\,\Psi(x_1), \Psi(x_0)\,\big] \in \operatorname{Hom}\big(\,E_{x_0}, E_{x_1}\,\big) \cong E_{x_1} \otimes E_{x_0}^*,$$

where we recall that $\operatorname{Cov}\big[\,\Psi(x_1), \Psi(x_0)\,\big]$ denotes the covariance operator of the jointly Gaussian random vectors $\Psi(x_1), \Psi(x_0)$.

The distribution Γ is uniquely determined by the distribution of the process $\mathcal{E}_\Gamma$ which in turn is uniquely determined by the collection $\big(\mathcal{K}(x_1, x_0)\,\big)_{x_0, x_1 \in M}$. This collection can be conveniently encoded as an *integration kernel*.

Consider the product $M \times M$ with its two canonical projections

$$M \overset{\pi_1}{\leftarrow} M \times M \overset{\pi_0}{\rightarrow} M, \;\; x_1 \leftarrow (x_1, x_0) \rightarrow x_0.$$

c Form the bundle

$$E \boxtimes E^* = \pi_1^* E \otimes \pi_0^* E^*.$$

Note that

$$\big(\,E \boxtimes E^*\,\big)_{(x_1, x_0)} = E_{x_1} \otimes E_{x_0}^* \cong \operatorname{Hom}\big(\,E_{x_0}, E_{x_1}\,\big).$$

Then $\mathcal{K}$ is intrinsically a section of $E \boxtimes E^*$. It is k-times differentiable and defines an integral operator

$$[\mathcal{K}] : L^2(E) \to L^2(E), \;\; [\mathcal{K}]u(x) = \int_M \mathcal{K}(x, y) u(y) \operatorname{vol}_g\big[\,dy\,\big]$$

Arguing as in the proof of Proposition 1.1.60 we deduce that this is a symmetric, nonnegative definite operator. The Karhune-Loève expansion continues to hold in this case as well and we deduce that any Gaussian C^ℓ-section of E can be described as a random series of C^ℓ-sections with coefficients independent random normal variables. Often in our applications $\mathcal{K}$ is the kernel of a smoothing operator.

Proposition 1.1.57 implies the following result.

Proposition 1.2.38. *Suppose that M is a smooth compact manifold, $E \to M$ is a smooth real vector bundle and (g, h, ∇) is a geometric structure on E. Recall that g is a Riemann* *metric on M, h is a metric on E and ∇ is a connection on E compatible with h. Fix $k \in \mathbb{N}_0$. Suppose that $(\psi_n)_{n\geq 1}$ is a sequence in $C^k(E)$ such that*

$$\sum_{n\geq 1} \|\psi_n\|_{C^k(E)} < \infty.$$

If $(A_n)_{n\geq 1}$ is a sequence of independent standard normal random variables, then the random series

$$\sum_{n\geq 1} A_n\psi_k$$

defines a Gaussian C^k-section of E. Its covariance kernel is

$$\mathcal{K}(x,y) = \sum_{n\geq 1} \psi_n(x) \otimes \psi_n(y)^{\dagger}, \quad x, y \in M$$

where $\psi_n(x) \otimes \psi_n(y)^{\dagger} : E_y \to E_x$ is the linear map

$$E_y \ni v \mapsto h\big(v, \psi_n(y)\big)\psi_n(x) \in E_x.$$

□

Example 1.2.39. Any centered Gaussian measure Γ on $C^k(E)$ tautologically defines a centered Gaussian random section of E given by the identity map $\mathbb{1} : C^k(E) \to C^k(E)$.

I want to point out a rather confusing fact. A fixed (deterministic) section ψ of E can also be viewed as a random section once we fix a Gaussian measure on $C^k(E)$. There will be arguments that will require juggling these two points of view. □

Example 1.2.40 (The main example). Suppose that (M, g) is a Riemann manifold. For $k \in \mathbb{N}$ We have a continuous linear map

$$C^k(M) \ni f \mapsto df \in C^{k-1}\big(T^*M\big).$$

If $F : M \to \mathbb{R}$ is a C^k Gaussian random function on M, then it induces a Gaussian measure $\mathbf{\Gamma}$ on $C^k(M)$ The pushforward of theis measure via the continuous linear map $d : C^k(M) \to C^{k-1}\big(T^*M\big)$ induces a Gaussian measure on $C^{k-1}\big(T^*M\big)$. Thus, the differential of a Gaussian random function on M is a Gaussian random section of T^*M. The main goal of this book is the investigation of the zero sets of such random differentials. □

Example 1.2.41. Let $E \to M$ be a smooth vector bundle equipped with a metric and a connection compatible with this metric. Suppose that $V \subset C^k(E)$ is a finite dimensional space of C^k-sections of E, $\psi_1, \dots, \psi_N$ is a basis of V and $X_1, \dots, X_N$ are independent standard normal random variables defined on the probability space $(\Omega, \mathcal{S}, \mathbb{P})$. Then the random linear combination

$$\Psi = \sum_{j=1}^{N} X_j\psi_j$$

is a centered Gaussian C^k-section. To see this consider the maps

$$\vec{X} : \Omega \to V, \;\; \omega \mapsto V, \;\; \Omega \ni \omega \mapsto \sum_j X_j(\omega)\psi_j \in V$$

Then $\Psi = i_V \circ \vec{X}$, where $i_V : V \hookrightarrow C^k(E)$ is the canonical inclusion. It is the composition of measurable maps and it is obviously centered Gaussian. Its covariance kernel is

$$\mathcal{K}^\Psi(x_1, x_0) = \sum_j \psi_j(x_1) \otimes \psi_j(x_0) \in E_{x_1} \otimes E_{x_0}$$

For each $x \in M$ we have a map

$$A_x : \mathbb{R}^N \to E_x, \;\; \mathbb{R}^N \ni \boldsymbol{u} \mapsto \sum_j u_j \psi_j(x).$$

Then

$$\mathcal{K}^\Psi(x, x) = A_x A_x^* \in \operatorname{End}(E_x).$$

We see that Ψ is ample iff A_x is onto, $\forall x \in M$.

More invariantly, note that for any $x \in M$ we have an evaluation map $\mathbf{Ev}_x : V \to E_x$, $V \ni v \mapsto v(x) \in E_x$. The Gaussian section Ψ is 0-ample iff these evaluation maps are onto, $\forall x \in M$. Algebraic geometers would say that the space of sections V is *ample.*

There is a more invariant way of describing this example. Fix an inner product on V so V becomes a Euclidean space. Let Γ_V the canonical Gaussian measure on this Euclidean space. Then $i_V : (\Gamma_V) \to C^k(E)$ is a random section. If $\psi_1, \dots, \psi_N$ is an orthonormal basis of V, then for any $\psi \in V$ we have

$$\Psi = i_V(\psi) = \sum_{j=1}^n X_j(\psi)\psi_j, \;\; X_j(\psi) = \big(\, \psi, \psi_j \,\big)_V.$$

Foe every $x \in M$ we have an evaluation map $\mathbf{Ev}_x : V \to E_x$ and the variance operator of $\Psi(x)$ is $\operatorname{Var}\big[\, \Psi(x) \,\big] = \mathbf{Ev}_x \, \mathbf{Ev}_x^*$. We see that Ψ is ample if $\mathbf{Ev}_x : V \to E_x$ is onto for all $x \in M$.

□

1.2.7. The differential geometry of a Gaussian ensemble. Let M be a smooth connected m-dimensional manifold and $E \to M$ a rank r smooth real vector bundle over M. Following [**114**], we will show that a smooth Gaussian random section of E canonically defines a metric on E and a connection compatible with the metric. Additionally, we will provide a probabilistic interpretation of this connection and its curvature.

A section $C \in C^k\big(\, E \boxtimes E \,\big)$ defines a family of bilinear maps

$$C_{\boldsymbol{p},\boldsymbol{q}} : E_{\boldsymbol{p}}^* \times E_{\boldsymbol{q}}^* \to \mathbb{R}, \;\; \boldsymbol{p}, \boldsymbol{q} \in M,$$

since $(E \boxtimes E)_{(\boldsymbol{p},\boldsymbol{q})} = E_{\boldsymbol{p}} \otimes E_{\boldsymbol{q}} \cong \big(\, E_{\boldsymbol{p}}^* \otimes E_{\boldsymbol{q}}^* \,\big)^*$. Such a section is called *symmetric* if for any $\boldsymbol{p}, \boldsymbol{q} \in M$ and any $\xi \in E_{\boldsymbol{p}}^*$, $\eta \in E_{\boldsymbol{q}}^*$ we have

$$C_{\boldsymbol{p},\boldsymbol{q}}(\xi, \eta) = C_{\boldsymbol{q},\boldsymbol{p}}(\eta, \xi).$$

Definition 1.2.42. A C^k-*correlator* on E is a symmetric section $C \in C^k\big(\, E \boxtimes E \,\big)$ such that $C_{\boldsymbol{p},\boldsymbol{p}}$ is positive definite for any $\boldsymbol{p} \in M$. □

Example 1.2.43. (a) Suppose that M is a properly embedded submanifold of the Euclidean space $\boldsymbol{U}$. Then the inner product $(-,-)_{\boldsymbol{U}}$ on $\boldsymbol{U}$ induces a correlator $C \in C^\infty(T^*M \boxtimes T^*M)$

defined by the equalities

$$C_{\boldsymbol{x},\boldsymbol{y}}(X,Y)=(X,Y)_{\boldsymbol{U}},\;\;\forall \boldsymbol{x},\boldsymbol{y}\in M,\;\; X\in T_{\boldsymbol{x}}M\subset \boldsymbol{U},\;\; Y\in T_{\boldsymbol{y}}M\subset\boldsymbol{U}.$$

(b) Let Ψ be an ample, centered Gaussian C^2-section of E. The random section Ψ defines a covariance *form*

$$C^{\Psi}\in C^k\big(\,E\boxtimes E\,\big),\quad C^{\Psi}(\boldsymbol{p},\boldsymbol{q})=C_{\Psi(\boldsymbol{p}),\Psi(\boldsymbol{q})},$$

where $C_{\Psi(\boldsymbol{p}),\Psi(\boldsymbol{q})}$ is the covariance form of the jointly Gaussian vectors $\Psi(\boldsymbol{p}),\Psi(\boldsymbol{q})$. Clearly, $C^{\Psi}(\boldsymbol{p},\boldsymbol{q})$ is symmetric. Since Ψ is ample, for any $\boldsymbol{p}\in M$, the Gaussian vector $\Psi(\boldsymbol{p})$ is nondegenerate and and thus its variance $C^{\Psi}_{\boldsymbol{p},\boldsymbol{p}}$ is positive definite. Hence C^{Ψ} is a correlator on E. □

Definition 1.2.44. A correlator $C\in C^k\big(\,E\boxtimes E)$ is called *stochastic* if it is the covariance form of an ample Gaussian C^k-section of E. □

Let $C\in C^k(E\boxtimes E)$ be a correlator where $k\geq 1$. By definition, it induces a metric on E^* and thus, by duality, a metric on E. We will denote these metrics by $(-,-)_{E^*,C}$ and respectively $(-,-)_{E,C}$. When no confusion is possible we will drop the subscript E or E^* from the notation. To simplify the presentation we adhere to the following conventions.

(i) We will use the Latin letters i,j,k to denote indices in the range $1,\dots,m=\dim M$.

(ii) We will use Greek letters α,β,γ to denote indices in the range $1,\dots,r=\operatorname{rank}(E)$.

Using the metric $(-,-)_C$ we can identify $C_{\boldsymbol{x},\boldsymbol{y}}\in E_{\boldsymbol{x}}\otimes E_{\boldsymbol{y}}$ with an element of

$$T_{\boldsymbol{x},\boldsymbol{y}}\in E_{\boldsymbol{x}}\otimes E^*_{\boldsymbol{y}}\cong \operatorname{Hom}(E_{\boldsymbol{y}},E_{\boldsymbol{x}}).$$

We will refer to $T_{\boldsymbol{x},\boldsymbol{y}}$ as the *tunneling map* from $E_{\boldsymbol{y}}$ to $E_{\boldsymbol{x}}$ associated to the correlator C. Note that $T_{\boldsymbol{x},\boldsymbol{x}}=\mathbb{1}_{E_{\boldsymbol{x}}}$. If we denote by $T^*_{\boldsymbol{x},\boldsymbol{y}}\in\operatorname{Hom}(E_{\boldsymbol{y}},E_{\boldsymbol{x}})$ the adjoint of $T_{\boldsymbol{x},\boldsymbol{y}}$ with respect to the metric $(-,-)_{E,C}$, then the symmetry of C implies that

$$T_{\boldsymbol{y},\boldsymbol{x}}=T^*_{\boldsymbol{x},\boldsymbol{y}}.$$

Lemma 1.2.45. *Fix a point $\boldsymbol{p}_0\in M$ and local coordinates $(x^i)_{1\leq i\leq m}$ in a neighborhood $\mathcal{O}$ of $\boldsymbol{p}_0$ in M. Suppose that $\underline{\boldsymbol{e}}(x)=(\boldsymbol{e}_\alpha(x))_{1\leq\alpha\leq r}$ is a local $(-,-)_C$-orthononomal frame of $E|_{\mathcal{O}}$. We regard it as an isomorphism of metric bundles $\boldsymbol{e}:\underline{\mathbb{R}}^r_{\mathcal{O}}\to E|_{\mathcal{O}}$ where $\underline{\mathbb{R}}^r_{\mathcal{O}}$ denotes the trivial bundle over $\mathcal{O}$ with fiber $\mathbb{R}^r$,*

$$\underline{\mathbb{R}}^r_{\mathcal{O}}:\;\;\big(\mathbb{R}^r\times\mathcal{O}\to\mathcal{O}\big).$$

. We obtain a smooth map

$$T(\underline{\boldsymbol{e}}):\mathcal{O}\times\mathcal{O}\to\operatorname{Hom}(\mathbb{R}^r),\;\;(x,y)\mapsto T(\underline{\boldsymbol{e}})_{x,y}=\underline{\boldsymbol{e}}(x)^{-1}T_{x,y}\underline{\boldsymbol{e}}(y).$$

Equivalently, $T(\underline{\boldsymbol{e}})_{x,y}$ makes commutative the diagram below.

$$\begin{array}{ccc}\mathbb{R}^r & \xleftarrow{T(\underline{\boldsymbol{e}})_{x,y}} & \mathbb{R}^r\\ {\scriptstyle \underline{\boldsymbol{e}}(x)}\downarrow & & \downarrow{\scriptstyle \underline{\boldsymbol{e}}(y)}\\ E_x & \xleftarrow[T_{x,y}]{} & E_y\end{array}\tag{1.2.34}$$

Then, for any $i=1,\dots,m$, the operator

$$\partial_{x^i}T(\underline{\boldsymbol{e}})_{x,y}|_{x=y}:\underline{\mathbb{R}}^r_y\to\underline{\mathbb{R}}^r_y,$$

is skew-symmetric.

Proof. We identify $\mathcal{O}\times\mathcal{O}$ with an open neighborhood of $(0,0)\in\mathbb{R}\times\mathbb{R}$ with coordinates (x^i,y^j). Introduce new coordinates $z^i:=x^i-y^i$, $s^j:=x^j+y^j$, so that $\partial_{x^i}=\partial_{z^i}+\partial_{s^i}$. We view the map $T(\underline{\boldsymbol{e}})$ as depending on the variables z,s. Note that

$$T(\underline{\boldsymbol{e}})_{0,s}=\mathbb{1},\ \ T(\underline{\boldsymbol{e}})_{-z,s}=T(\underline{\boldsymbol{e}})^*_{z,s},\ \ \forall z,s.$$

We deduce that

$$\partial_{s^i}T(\underline{\boldsymbol{e}})|_{0,s}=\partial_{s^i}T(\underline{\boldsymbol{e}})|^*_{0,s}=0,$$

$$\partial_{x^i}T(\underline{\boldsymbol{e}})|_{0,s}=\partial_{z^i}T(\underline{\boldsymbol{e}})|_{0,s}+\partial_{s^i}T(\boldsymbol{e})|_{0,s}=\partial_{z^i}T(\underline{\boldsymbol{e}})|_{0,s},$$

$$\big(\,\partial_{x^i}T(\underline{\boldsymbol{e}})|_{0,s}\,\big)^*=\partial_{x^i}T(\underline{\boldsymbol{e}})^*|_{0,s}=-\partial_{z_i}T(\underline{\boldsymbol{e}})|_{0,s}+\partial_{s^i}T(\underline{\boldsymbol{e}})|_{0,s}=-\partial_{x^i}T(\underline{\boldsymbol{e}})|_{0,s}.$$

□

Given a coordinate neighborhood with coordinates (x^i) and a local isomorphism of metric vector bundles (local orthonormal frame) $\underline{\boldsymbol{e}}:\underline{\mathbb{R}}^r_{\mathcal{O}}\to E|_{\mathcal{O}}$ as above, we define the skew-symmetric endomorphisms

$$\Gamma_i(\underline{\boldsymbol{e}}):\underline{\mathbb{R}}^r_{\mathcal{O}}\to\underline{\mathbb{R}}^r_{\mathcal{O}},\ \ i=1,\dots,m=\dim M,\ \ \Gamma_i(\underline{\boldsymbol{e}})_y=-\partial_{x^i}T_{x,y}|_{x=y}. \tag{1.2.35}$$

We obtain a 1-form with matrix coefficients $\Gamma(\underline{\boldsymbol{e}}):=\sum_i\Gamma_i(\underline{\boldsymbol{e}})dy^i$. The operator

$$\nabla^{\underline{\boldsymbol{e}}}=d+\Gamma(\underline{\boldsymbol{e}}) \tag{1.2.36}$$

is then a connection on $\underline{\mathbb{R}}^r_{\mathcal{O}}$ compatible with the natural metric on this trivial bundle. The isomorphism $\underline{\boldsymbol{e}}$ induces a metric connection $\underline{\boldsymbol{e}}_*\nabla^{\underline{\boldsymbol{e}}}$ on $E|_{\mathcal{O}}$.

Suppose that $\underline{\boldsymbol{f}}:\underline{\mathbb{R}}^r_{\mathcal{O}}\to E|_{\mathcal{O}}$ is another orthonormal frame of $E_{\mathcal{O}}$ related to $\underline{\boldsymbol{e}}$ via a transition map

$$g:\mathcal{O}\to O(r),\ \ \underline{\boldsymbol{f}}=\underline{\boldsymbol{e}}\cdot g.$$

Then

$$T(\underline{\boldsymbol{f}})_{x,y}=g^{-1}(x)T(\underline{\boldsymbol{e}})_{x,y}g(y).$$

We denote by d_x the differential with respect to the x variable. We deduce

$$\Gamma(\underline{\boldsymbol{f}})_y=-d_xT(\underline{\boldsymbol{f}})_{x,y}|_{x=y}$$

$$=-\big(\,d_xg^{-1}(x)\,\big)_{x=y}\cdot\underbrace{T(\underline{\boldsymbol{e}})_{y,y}}_{=\mathbb{1}}\cdot g(y)-g^{-1}(y)\big(\,d_xT(\underline{\boldsymbol{e}})_{x,y}\big)|_{x=y}g(y)$$

$$=g^{-1}(y)dg(y)g^{-1}(y)\cdot g(y)+g^{-1}(y)\Gamma(\underline{\boldsymbol{e}})_yg(y)=g(y)^{-1}dg(y)+g^{-1}(y)\Gamma(\underline{\boldsymbol{e}})_yg(y).$$

Thus

$$\Gamma(\underline{\boldsymbol{e}}\cdot g)=g^{-1}dg+g^{-1}\Gamma(\underline{\boldsymbol{e}})g.$$

This shows that for any local orthonormal frames $\underline{\boldsymbol{e}}$, $\underline{\boldsymbol{f}}$ of $E|_{\mathcal{O}}$ we have

$$\underline{\boldsymbol{e}}_*\nabla^{\underline{\boldsymbol{e}}}=\underline{\boldsymbol{f}}_*\nabla^{\underline{\boldsymbol{f}}}.$$

We have thus proved the following result.

Proposition 1.2.46. *If $E \to M$ is a smooth real vector bundle, then any correlator C on M induces a canonical metric $(-,-)_C$ on E and a connection ∇^C compatible with this metric. More explicitly, if $\mathcal{O} \subset M$ is an coordinate neighborhood on M and $\underline{\boldsymbol{e}} : \mathbb{R}^r_{\mathcal{O}} \to E|_{\mathcal{O}}$ is an orthogonal trivialization , then ∇^C is described by*

$$\nabla^C = d + \sum_i \Gamma_i(\underline{\boldsymbol{e}}) dx^i,$$

where the skew-symmetric $r \times r$-matrix $\Gamma_i(\underline{\boldsymbol{e}})$ is given by (1.2.35). We will refer to ∇^C as the correlator connection. □

Remark 1.2.47. Suppose that we fix local coordinates (x^i) near a point $\boldsymbol{p}_0$ such that $x^i(\boldsymbol{p}_0) = 0$. We denote by $P_{x,0}$ the parallel transport of ∇^C from 0 to X along the line segment from 0 to x. Then

$$P_{0,0} = \mathbb{1}_{E_0} = T_{0,0}, \;\; \partial_{x^i} P_{x,0}|_{x=0} = -\Gamma_i(0) = \partial_{x^i,0} T_{x,0}|_{x=0}.$$

We see that the tunneling map $T_{x,0}$ is a first order approximation at 0 of the parallel transport map $P_{x,0}$ of the connection ∇^C. □

When the correlator is stochastic, this connection can be given a probabilistic description. Fix an ample Gaussian measure Γ on $C^k(E)$. Denote by $C \in C^k\big(E \boxtimes E\big)$ its correlator and by ∇^C the connection it determines on E. As we mentioned earlier, a section $\psi \in C^k(E)$ has a dual incarnation: a deterministic one, as a section of E and a probabilistic one, as an element of the probability space $\big(C^k(E), \Gamma\big)$.

Fix a point $\boldsymbol{p}_0$, a coordinate neighborhood $\mathcal{O}$ of $\boldsymbol{p}_0$ in M and orthonormal framings $\underline{\boldsymbol{e}} : \mathbb{R}^m_{\mathcal{O}} \to E|_{\mathcal{O}}$ as in Lemma 1.2.45. We get a random map $\Phi : \mathcal{O} \to \mathbb{R}^m$, $\Phi(x) = \underline{\boldsymbol{e}}(x)^{-1}\psi(y)$; see diagram (1.2.34). By definition, the covariance form of $\psi(x)$ is given by the metric on E_y. The map $\underline{\boldsymbol{e}}(x)$ is an isometry so that the variance operator of $\Phi(x)$ is $\mathbb{1}_{\mathbb{R}^m}$. Thus

$$\partial_{x^i} T(\underline{\boldsymbol{e}})_{x,y}|_{x=y} = R_{\partial_{x^i}\Phi(x), \Phi(x)}$$

where $R_{-,-}$ is the regression operator (1.1.16) . We deduce from the regression formula (1.1.17) and Proposition 1.2.46 that

$$\nabla^C \Phi(x) = d\Phi(x) - \mathbb{E}\big[\, d\Phi(x) \,\|\, \Phi(x)\,\big]. \tag{1.2.37}$$

In particular we deduce the following result.

Corollary 1.2.48. *For any $\psi \in C^k(E)$, and any $x \in M$, the random vector $\nabla^C\psi(x)$ is independent of the random vector $\psi(x)$.* □

In [**53**, Prop. 1.1.3] it is shown that there is only one connection ∇ on E, compatible with the metric induced by the correlator C such that, for any $x \in M$, the random vector $\nabla^C\psi(x)$ is independent of the random vector $\psi(x)$. The authors refer to this connection as the *LeJan-Watanabe* or *L-W connection.*

Proposition 1.2.49. *Suppose that C is a stochastic correlator on E defined by an ample Gaussian ensemble C^2 random sections of E. Denote by $\boldsymbol{u}$ a random section in this ensemble. Fix a point $\boldsymbol{p}_0$, local coordinates (x^i) on M near $\boldsymbol{p}_0$ such that $x^i(\boldsymbol{p}_0) = 0$ $\forall i$, and* *a local $(-,-)_C$-orthonormal frame $\big(\boldsymbol{e}_\alpha(x)\big)_{1\leq\alpha\leq r}$ of E in a neighborhood of $\boldsymbol{p}_0$ which is is synchronous at $\boldsymbol{p}_0$,*

$$\nabla^C \boldsymbol{e}_\alpha|_{\boldsymbol{p}_0} = 0, \;\; \forall\alpha.$$

Denote by F the curvature of ∇^C,

$$F = \sum_{ij} F_{ij}(x)dx^i\wedge dx^j, \;\; F_{ij}(x)\in \operatorname{End}(E_{\boldsymbol{p}_0}).$$

Then $F_{ij}(0)$ is the endomorphism of $E_{\boldsymbol{p}_0}$ which in the frame $\boldsymbol{e}_\alpha(\boldsymbol{p}_0)$ is described by the $r\times r$ matrix with entries

$$F_{\alpha\beta|ij}(0) := \mathbb{E}\big[\,\partial_{x^i}u_\alpha(x)\partial_{x^j}u_\beta(x)\,\big]|_{x=0} - \mathbb{E}\big[\,\partial_{x^j}u_\alpha(x)\partial_{x^i}u_\beta(x)\,\big]|_{x=0}, \;\; 1\leq\alpha,\beta\leq r, \tag{1.2.38}$$

where $u_\alpha(x)$ is the random function

$$u_\alpha(x) := \big(\,\boldsymbol{u}(x), \boldsymbol{e}_\alpha(x)\,\big)_C.$$

Proof. The random section $\boldsymbol{u}$ has the local description

$$\boldsymbol{u} = \sum_\alpha u_\alpha(x)\boldsymbol{e}_\alpha(x).$$

Then $T(x,y)$ is a linear map $E_y\to E_x$ given by the $r\times r$ matrix

$$T(x,y) = \big(\,T_{\alpha\beta}(x,y)\,\big)_{1\leq\alpha,\beta\leq r}, \;\; T_{\alpha\beta}(x,y) = \mathbb{E}\big[\,u_\alpha(x)u_\beta(y)\,\big].$$

The coefficients of the connection 1-form $\Gamma = \sum_i \Gamma_i dx^i$ are endomorphisms of E_x given by $r\times r$ matrices $\Gamma_i(x) = \big(\,\Gamma_{\alpha\beta|i}(x)\,\big)_{1\leq\alpha,\beta\leq r}$. More precisely, we have

$$\Gamma_{\alpha\beta|i}(x) = -\mathbb{E}\big[\,\partial_{x^i}u_\alpha(x)u_\beta(x)\,\big]. \tag{1.2.39}$$

Because the frame $\big(\,\boldsymbol{e}_\alpha(x)\,\big)$ is synchronous at $x=0$ we deduce that, *at* $\boldsymbol{p}_0$, we have $\Gamma_i(0)=0$ and

$$F(\boldsymbol{p}_0) = \sum_{i<j} F_{ij}(x)dx^i\wedge dx^j \in \operatorname{End}(E_{\boldsymbol{p}_0})\otimes\Lambda^2 T^*_{\boldsymbol{p}_0}M, \;\; F_{ij} = \partial_{x^i}\Gamma_j(\boldsymbol{p}_0) - \partial_{x^j}\Gamma_i(\boldsymbol{p}_0).$$

The coefficients $F_{ij}(x)$ are $r\times r$ matrices with entries $F_{\alpha\beta|ij}(x)$, $1\leq\alpha,\beta\leq r$. Moreover,

$$\begin{aligned}
F_{\alpha\beta|ij}(0) &= \partial_{x^j}\Gamma_{\alpha\beta|j}(0) - \partial_{x^j}\Gamma_{\alpha\beta|i}(0)\\
&\stackrel{(1.2.39)}{=} \partial_{x^j}\mathbb{E}\big[\,\partial_{x^i}u_\alpha(x)u_\beta(x)\,\big)|_{x=0} - \partial_{x^i}\boldsymbol{E}\big(\,\partial_{x^j}u_\alpha(x)u_\beta(x)\,\big)|_{x=0}\\
&= \mathbb{E}\big(\,\partial^2_{x^jx^i}u_\alpha(x)u_\beta(x)\,\big]|_{x=0} + \mathbb{E}\big[\,\partial_{x^i}u_\alpha(x)\partial_{x^j}u_\beta(x)\,\big]|_{x=0}\\
&\quad -\mathbb{E}\big[\,\partial^2_{x^ix^j}u_\alpha(x)u_\beta(x)\,\big]|_{x=0} - \mathbb{E}\big[\,\partial_{x^j}u_\alpha(x)\partial_{x^i}u_\beta(x)\,\big]|_{x=0}\\
&= \mathbb{E}\big[\,\partial_{x^i}u_\alpha(x)\partial_{x^j}u_\beta(x)\,\big]|_{x=0} - \mathbb{E}\big(\,\partial_{x^j}u_\alpha(x)\partial_{x^i}u_\beta(x)\,\big]|_{x=0}.
\end{aligned}$$

□

Example 1.2.50. Suppose that $\Phi : M \to \mathbb{R}$ is a Gaussian C^3-function on the smooth manifold M such that the differential $d\Phi$ is an ample Gaussian C^2-section of T^*M. The correlator of $d\Phi$ defines a metric on TM and a connection compatible with it. This turns out to be the Levi-Civita connection of the correlator metric; see [**117**, Sec. 4.2.5]. For an alternate description of this connection we refer to [**1**, Sec.12.2].

As a special case, suppose that M is a compact smooth submanifold of the Euclidean space $\boldsymbol{U}$. Denote by $(-,-)$ the inner product on $\boldsymbol{U}$ and by Γ the canonical Gaussian measure on $\boldsymbol{U}$. We obtain a Gaussian function on M,

$$\Phi : \boldsymbol{U} \times M \to \mathbb{R}, \;\; \boldsymbol{U} \times M \ni (\boldsymbol{u}, \boldsymbol{x}) \mapsto \Phi_{\boldsymbol{u}}(\boldsymbol{x}) = (\boldsymbol{u}, \boldsymbol{x}).$$

The differential $d\Phi$ is an ample[8] Gaussian section of T^*M and the correlator metric on TM is the induced metric on M.

In this case the curvature formula (1.2.38) implies Theorema Egregium stating that the curvature is an intrisic invariant of the submanifold. The classical approach to Theorema Egregium goes through the second fundamental form of M. The probabilistic approach bypasses this object. However, the second fundamental form has many other fundamendal uses. For details I refer to [**117**, Sec. 4.2.5]. □

[8]Can you see why?

Chapter 2

The Gaussian Kac-Rice formula

Suppose that $\boldsymbol{U}$ and $\boldsymbol{V}$ are two Euclidean spaces of the same dimension, $\mathscr{V} \subset \boldsymbol{V}$ an open subset of $\boldsymbol{V}$, and $\Phi : \mathscr{V} \to \boldsymbol{U}$ a centered Gaussian map that is a.s. C^k, with k to be specified later. "Typically", the zero set of Φ is discrete so that for any compact subset $K \subset \mathscr{V}$ the set $\{\Phi = 0\} \cap K$ is finite. We denote by Z_K or $Z[K, \Phi]$ its cardinality. In this section we investigate the basic invariants of this random variable: expectation, variance and higher momentums. The Kac-Rice formula is essentially a description of these invariants as integrals of certain densities over $\mathscr{V}$. However, before we state and prove this formula there are a few technical but important issues to address.

2.1. Generic transversality

Suppose $\boldsymbol{U}$ and $\boldsymbol{V}$ are two real Euclidean spaces of dimensions

$$d = \dim \boldsymbol{U} \leq D := \dim \boldsymbol{V}.$$

Let $\mathscr{V} \subset \boldsymbol{V}$ be an open set.

Definition 2.1.1. Suppose that $X : \Omega \times \mathscr{V} \to \boldsymbol{U}$ is a C^k random field. We say that X satisfies the *standard conventions* if

- the probability space $(\Omega, \mathcal{S}, \mathbb{P})$ is $\mathbb{P}$-complete, and
- For any $0 \leq j \leq k$, the j-th differential

  $$D^j X : \Omega \times \mathscr{V} \to \mathbf{Sym}^j \left(\boldsymbol{V}, \boldsymbol{U} \right)$$

  is $\mathcal{S} \otimes \mathcal{B}_{\mathscr{V}}$-measurable and separable; see Definition 1.2.3. Above $\mathbf{Sym}^j(\boldsymbol{V}, \boldsymbol{U})$ denotes the space of symmetric j-linear maps $\boldsymbol{V}^j \to \boldsymbol{U}$.

□

For example, if Γ is a Gaussian measure on $\Omega = C^k(\mathscr{V}, \boldsymbol{U})$, and $\mathcal{S}$ is the Γ-completion of the Borel sigma-algebra of Ω, then the resulting random field $\mathbf{Ev}^\Gamma : \Omega \times \mathscr{V} \to \boldsymbol{U}$ satisfies the standard conventions; see Example 1.2.18.

☞ *In the sequence will tacitly assume that the random fields satisfy the standard conventions.*

Let me recall a classical transversality result frequently used in differential topology [**71**, Chap.3]. The origin $0 \in \boldsymbol{U}$ is a regular value of most $F \in C^\infty(\mathscr{V}, \boldsymbol{U})$ and thus for a "typical" F the level set $F^{-1}(0)$ is a submanifold of codimension $d = \dim \boldsymbol{U}$ and we do not expect it to intersect a submanifold of $\mathscr{V}$ of dimension $< d$. The next result is a quantitative version of this fact.

Lemma 2.1.2 (Bulinskaya)**.** *Suppose that*

$$X : \big(\,\Omega, \mathcal{S}, \mathbb{P}\big) \times \mathcal{V} \to \boldsymbol{U},\ \ \Omega \times \mathscr{V} \ni (\omega, v) \mapsto X_\omega(v) \in \boldsymbol{U}$$

is an a.s. C^1 *Gaussian random field. Assume that X is ample, i.e.,*

for any $v \in \mathscr{V}$ the Gaussian vector

$$\Omega \ni \omega \mapsto X_\omega(v) \in \boldsymbol{U} \tag{A_0}$$

is nondegenerate.

Fix $u_0 \in \boldsymbol{U}$ and let $K \subset \mathscr{V}$ be a compact set of Hausdorff dimension $< d = \dim \boldsymbol{U}$. Then the set[1]

$$A := \big\{\, \omega \in \Omega;\ \ \exists v \in K \ \ \textit{such that}\ X(v) = u_0 \,\big\} \tag{2.1.1}$$

is negligible.

Proof. I follow the argument in the proof of [**1**, Lemma 11.2.10]. Denote by X' the differential of X, by $\|-\|$ the Euclidean norms on $\boldsymbol{U}$ and $\boldsymbol{V}$, and by $\|-\|_{\mathrm{op}}$ the operator norm on $\mathrm{Hom}(\boldsymbol{V}, \boldsymbol{U})$. Let

$$C_\omega(v) := \|X_\omega\| + \|X'_\omega(v)\|_{\mathrm{op}}.$$

For every compact set $S \subset \mathcal{V}$ we set

$$C_\omega(S) := \sup_{v \in S} C_\omega(v).$$

Fernique's inequality (1.1.32) shows that $C(S) \in L^1$ so $\mathbb{P}\big[\,C(S) < \infty\,\big] = 1$. Hence, for every $\varepsilon > 0$, there exists $M_\varepsilon = M_\varepsilon(S) > 0$ such that

$$\mathbb{P}\big[\,C(S) < M_\varepsilon\,\big] > 1 - \varepsilon. \tag{2.1.2}$$

Choose S to be a closed ball of radius $r > 0$ centered at v_0 and contained in $\mathcal{V}$ and set $C(v_0, r) := C(S)$. We deduce from the mean value theorem that

$$\|\,X_\omega(v) - X_\omega(v_0)\| \le C_\omega(v_0, r) \cdot r.$$

---

[1] The measurability of A is tricky. Consider the space $\mathfrak{X} := \Omega \times \mathscr{V}$ equipped with the product σ-algebra. The map $\mathfrak{X} \ni (\omega, v) \overset{\Phi}{\mapsto} (X_\omega(v), v) \in \boldsymbol{U} \times \mathscr{V}$ is measurable as composition of measurable maps

$$\Omega \times \mathcal{V} \overset{X \times \mathbb{1}}{\longrightarrow} C^1(\mathscr{V}, U) \times \mathscr{V} \overset{\mathbf{Ev} \times \mathbb{1}}{\longrightarrow} \boldsymbol{U} \times \mathscr{V}.$$

The subset $\mathcal{Z} := \Phi^{-1}(\{u_0\} \times K)$ is measurable in $\mathfrak{X}$. If we denote by π the natural projection $\mathfrak{X} \to \Omega$ then $A = \pi(\mathcal{Z})$ and according to [**36**, Prop. 8.4.4] it is $\mathcal{S}$-measurable if $\mathcal{S}$ is $\mathbb{P}$-complete.

For $r < \operatorname{dist}(K, \partial\mathcal{V})$ we set

$$C_\omega(K,r) := \sup_{v_0 \in K} C_\omega(v_0, r) \leq C_\omega(K_r), \quad K_r := \{ v \in \boldsymbol{V};\ \operatorname{dist}(v, K) \leq r \},$$

and

$$\operatorname{osc}_\omega(r) := \sup_{\substack{v_1, v_2 \in K, \\ \|v_1 - v_2\| \leq r}} \|X_\omega(v_1) - X_\omega(v_2)\|.$$

Note that

$$\operatorname{osc}_\omega(r) \leq C_\omega(K,r) r.$$

Consider the event

$$E_\varepsilon(r) := \big\{ \operatorname{osc}_\omega(r) \leq M_\varepsilon(K_r)\, r \big\}.$$

We set $M_\varepsilon(r) := M_\varepsilon(K_r)$. We deduce from (2.1.2) that

$$\mathbb{P}\big[\, E_\varepsilon(r) \,\big] > 1 - \varepsilon.$$

Pick a sequence $\hbar_n \searrow 0$. Since K has Hausdorff dimension $< d$, its d-dimensional Hausdorff measure is zero, and we deduce that there exists a sequence of radii $r_n \searrow 0$ and, for any n, there exists a finite collection of closed balls $(B_{n,j})_{j \in J_n}$, of radii $r_{n,j} < r_n$, covering K, such that

$$\sum_{j \in J_n} \big(r_{n,j} \big)^d \leq \hbar_n.$$

Set

$$A_{n,j} := \big\{ \omega \in \Omega;\ \ \exists v \in K \cap B_{n,j} \ \text{ such that } X_\omega(v) = u_0 \big\} \subset A, .$$

where A is defined as (2.1.1). Fix $\varepsilon > 0$ and $r > 0$ sufficiently small. Then

$$\mathbb{P}\big[\, A \,\big] \leq \sum_j \mathbb{P}\big[\, A_{n,j} \cap E_\varepsilon(r_n) \,\big] + \mathbb{P}\big[\, E_\varepsilon(r_n)^c \,\big] \leq \sum_j \mathbb{P}\big[\, A_{n,j} \cap E_\varepsilon(r_n) \,\big] + \varepsilon. \tag{2.1.3}$$

Denote by $v_{n,j}$ the center of $B_{n,j}$. Observe that $A_{n,j} \neq \emptyset$ iff there exists v such that $\|v - v_{n_j}\| \leq r_{n,j}$ and $X(v) = u_0$. On $E_\varepsilon(r_n)$ we have

$$\big\|\, X(v_{n,j}) - u_0 \,\big\| = \big\|\, X(v_{n,j}) - X(v) \,\big\| \leq M_\varepsilon(r_n)\, r_{n,j}.$$

This shows that

$$A_{n,j} \cap E_\varepsilon(r_n) \subset \Big\{ \big\|\, X(v_{n,j}) - u_0 \,\big\| < M_\varepsilon(r_n)\, r_{n,j} \Big\}.$$

Denote by $\boldsymbol{\omega}_d$ the volume of the unit d-dimensional Euclidean ball, by $p_{X(v)}$ the probability density of $X(v)$ and set

$$L := \sup_{v \in K_{r_1}} \sup_{u \in \boldsymbol{U}} p_{X(v)}(u).$$

The ampleness ssumption (A_0) implies $L < \infty$. We deduce

$$\mathbb{P}\big[\, \big\{ \big\|\, X(v_{n,j}) - u_0 \,\big\| < M_\varepsilon(r_n)\, r_{n,j} \big\} \,\big] \leq \underbrace{L \boldsymbol{\omega}_d M_\varepsilon(r_n)^d}_{=:\Xi_\varepsilon(r_n)} r_{n,j}^d,$$

and

$$\sum_j \mathbb{P}\big[\, A_{n,j} \cap E_\varepsilon(r_n) \,\big] \leq \Xi_\varepsilon(r_n) \sum_j r_{n,j}^d \leq \Xi_\varepsilon(r_n) \hbar_n \leq \Xi_\varepsilon(r_1) \hbar_n.$$

Now choose n such that $\Xi_\varepsilon(r_1)\hbar_n \leq \varepsilon$ to conclude from (2.1.3) that

$$\mathbb{P}\big[\, A \,\big] \leq 2\varepsilon, \ \ \forall \varepsilon > 0.$$

□

Theorem 2.1.3. *Suppose that* $(\Omega, \mathcal{S}, \mathbb{P})$ *is a probability space and*

$$X : \Omega \times \mathscr{V} \to \boldsymbol{U}, \;\; \Omega \times \mathcal{V} \ni (\omega, v) \mapsto X_\omega(v) \in \boldsymbol{U}$$

is Gaussian field with the following properties.

(i) *The random field* X *is* a.s. C^2.

(ii) *The Gaussian vector*

$$Y : \Omega \times \mathcal{V} \times \boldsymbol{U} \setminus \{0\} \to \boldsymbol{U} \times \boldsymbol{V}, \;\; (\omega, v, \dot{u}) \mapsto (X_\omega(v), X'_\omega(v)^*\dot{u})$$

nondegenerate for any $v \in \mathscr{V}$*. Above,* $X'_\omega(v)^* : \boldsymbol{U} \to \boldsymbol{V}$ *denotes the adjoint of the differential* $X'_\omega(v) : \boldsymbol{V} \to \boldsymbol{U}$ *of* X_ω *at* v*.*

Then $0 \in \boldsymbol{U}$ *is* a.s. *a regular value of* X*, i.e.,*

$$\mathbb{P}\big[\, \{\, \omega;\; 0 \text{ is a regular value of } X_\omega : \mathcal{V} \to \boldsymbol{U} \,\}\, \big] = 1.$$

Proof. Fix a closed ball $B \subset \mathcal{V}$ and denote by $S(\boldsymbol{U})$ the unit sphere in $\boldsymbol{U}$. Let us show that a.s., 0 is a regular value of $X|_B$. This means that for any solution $v \in B$ of $X = 0$ the adjoint of the differential $X'(v)$ is one-to-one, i.e., the equation

$$Y(v, \dot{u}) = 0 \Longleftrightarrow X(v) = 0, \;\; X'(v)^*\dot{u} = 0,$$

has no solution $(v, \dot{u}) \in B \times S(\boldsymbol{U})$. Since $\dim B \times S(\boldsymbol{U}) < \dim(\boldsymbol{U} \times \boldsymbol{V})$ we deduce from Bulinskaya's Lemma 2.1.2 that this happens a.s.

□

Theorem 2.1.3 can be substantially improved when $\dim \boldsymbol{U} = \dim \boldsymbol{V}$

Theorem 2.1.4. *Suppose that* $\dim \boldsymbol{V} = \dim \boldsymbol{U} = d$*,* $(\Omega, \mathcal{S}, \mathbb{P})$ *is a probability space and*

$$X : \Omega \times \mathcal{V} \to \boldsymbol{U}, \;\; \Omega \times \mathcal{V} \ni (\omega, v) \mapsto X_\omega(v) \in \boldsymbol{U}$$

is an ample C^1 *Gaussian field. Then* $0 \in \boldsymbol{U}$ *is* a.s. *a regular value of* X*.*

Proof. We follow the approach in [7, Sec.4]. Fix a closed ball $B \subset \mathcal{V}$. Consider the quantities

$$T = \liminf_{r \searrow 0} T_r, \;\; T_r(\omega) := \frac{1}{\boldsymbol{\omega}_d r^d} \mathcal{H}_d\big[\, \{ v \in B;\; \|X(v)\| \leq r \} \,\big],$$

where $\mathcal{H}_d$ denotes the d-dimensional Hausdorff measure on $\boldsymbol{V}$. In this case it coincides with the Lebesgue measure. Denote by J_v the Jacobian of the map X at v,

$$J_v = \sqrt{\det\big(\, X'(v) X'(v)^* \,\big)}.$$

Since $\boldsymbol{U}$ and $\boldsymbol{V}$ are Euclidean spaces of the same dimension we have $J_v = \big|\, \det X'(v) \,\big|$.

Since X is ample we deduce that the random variable T defined above is a.s. finite. We set

$$\mathcal{Z}^s := \big\{\, \exists v \in B, \;\; X(v) = 0, \;\; J_v = 0 \,\big\}.$$

We will show that $\mathbb{P}\big[\, \mathcal{Z}^s \neq \emptyset \,\big] = 0$. Set

$$M := \sup_{v \in B} \|X'(v)\|_{\mathrm{op}}, \;\; N(\varepsilon) := \sup_{v \in B,\, 0 < \|\dot{v}\| < \varepsilon} \frac{\|X(v_0 + \dot{v}) - X(v_0) - X'(v_0)\dot{v}\|}{\|\dot{v}\|}.$$

Both random variables M and $N(\varepsilon)$ are a.s. finite and $N(\varepsilon) \to 0$ a.s. as $\varepsilon \searrow 0$.

Let $v_0 \in \mathcal{Z}^s$. Lemma 2.1.2 shows that $v_0 \in \boldsymbol{int}\, B$ a.s.. Set

$$K_0 := \ker X'(v_0) \subset \boldsymbol{V}, \;\; k := \dim K_0^\perp.$$

Since $J_{v_0} = 0$ we deduce that $k < d = \dim \boldsymbol{V}$. Any vector $\dot{v} \in \boldsymbol{V}$ decomposes as

$$\dot{v} = \dot{v}_0 + \dot{v}^\perp, \;\; \dot{v}_0 \in K_0, \;\; \dot{v}^\perp \in K_0^\perp.$$

Then

$$\begin{aligned} \|X(v_0 + \dot{v})\| &\leq \|X(v_0 + \dot{v}_0 + \dot{v}^\perp) - X(v_0 + \dot{v}_0)\| + \|X(v_0 + \dot{v}_0)\| \\ &\leq M\|\dot{v}^\perp\| + \|\dot{v}_0\| N\big(\,\|\dot{v}_0\|\,\big). \end{aligned}$$

Let $\varepsilon > 0$ such that $N(\varepsilon) < 1$ and suppose that

$$\|\dot{v}_0\| \leq \varepsilon, \;\; \|\dot{v}^\perp\| \leq \varepsilon N(\varepsilon). \tag{2.1.4}$$

We deduce that

$$\|X(v_0 + \dot{v}_0 + \dot{v}^\perp)\| \leq r(\varepsilon) := (M+1)\varepsilon N(\varepsilon).$$

The polydisk

$$P_\varepsilon := \big\{\, v \in B; \;\; v = v_0 + \dot{v}, \;\; \dot{v} \text{ satisfies } (2.1.4) \,\big\}$$

is a.s. contained in B for $\varepsilon > 0$ sufficiently small. Thus

$$\begin{aligned} T_{r(\varepsilon)} &= \frac{1}{\boldsymbol{\omega}_d r(\varepsilon)^d} \mathcal{H}_d\big[\,\{v \in B; \; \|X(v)\| \leq r(\varepsilon)\}\,\big] \geq \frac{1}{\boldsymbol{\omega}_d r(\varepsilon)^d} \mathcal{H}_d\big[\, P_\varepsilon \,\big] \\ &= \frac{\mathit{const.} \times \varepsilon^d N(\varepsilon)^k}{\boldsymbol{\omega}_d \varepsilon^d N(\varepsilon)^d} = \mathit{const} N(\varepsilon)^{k-d} \to \infty \;\text{ as } \varepsilon \searrow 0. \end{aligned}$$

Hence

$$\mathcal{Z}^s \neq \emptyset \subset \{T = \infty\},$$

so $\mathbb{P}\big[\,\mathcal{Z}^s = \emptyset\,\big] = 1$. □

Remark 2.1.5. To better understand the idea behind the above proof it helps to have in mind the following elementary yet suggestive example. Consider the map

$$F : \mathbb{R}^2 \to \mathbb{R}^2, \;\; F(x,y) = (x, y^2).$$

Then

$$T_r := \big\{\|F\| \leq r\,\big\} = \big\{\, x^2 + y^4 \leq r^2 \,\big\} \supset S_r := \big\{\, |x| \leq 2^{-1/2} r, \;\; |y| \leq 2^{-1/4}\sqrt{r}\,\big\},$$

and $\mathcal{H}_2\,(S_r) = 2^{-3/4} r^{3/2}$. Hence

$$\frac{\mathcal{H}_2(T_r)}{\pi r^2} \geq 2^{-3/4} r^{-1/2} \nearrow \infty \;\text{ as } r \searrow 0.$$

□

Corollary 2.1.6. *Let $\mathscr{V}$ be an open subset of the Euclidean space $\boldsymbol{V}$. Suppose that $F : \mathscr{V} \to \mathbb{R}$ is a C^2 Gaussian function that such that its differential is ample, i.e., for any $v \in \mathscr{V}$ the Gaussian vector $dF(v) \in \boldsymbol{V}$ is nondegenerate. Denote by $\operatorname{Hess}_F(v)$ the Hessian of F at v. Then F is* a.s. *a Morse function, i.e.,*

$$\mathbb{P}\big[\,\{\,\exists v, \;\; df(v) = 0, \;\; \det \operatorname{Hess}_F(v) = 0\,\}\,\big] = 0.$$

□

Remark 2.1.7. Sard's transversality theorem requires a bit of regularity. Suppose that $F : \mathcal{V} \to \boldsymbol{U}$ is a C^k map. In [57, Thm. 3.4.3] it is shown that if $k > \dim \boldsymbol{V} - \dim \boldsymbol{U}$, then the set of critical values of F is negligible in $\boldsymbol{U}$. However, if $k \leq \dim \boldsymbol{V} - \dim \boldsymbol{U}$, then there exist C^k-maps $\mathcal{V} \to \boldsymbol{U}$ for which the set of critical values is not negligible in $\boldsymbol{U}$; see [57, Sec. 3.4.4].

In geometry the generic transversality is traditionally obtained as follows. Suppose that N is a positive integer and

$$F : \mathbb{R}^N \times \mathcal{V} \to \boldsymbol{U}, \;\; (\lambda, v) \to F_\lambda(v)$$

is a C^k-map, $k > \dim \boldsymbol{V} - \dim \boldsymbol{U}$. We view it as a family in $C^k(\mathcal{V}, \boldsymbol{U})$ parametrized by $\lambda \in \mathbb{R}^N$. We assume that the family is sufficiently large, i.e., satisfies the ampleness condition

$$0 \text{ is a regular value } F. \tag{$*$}$$

Then

$$Z = \big\{ (\lambda, v) \in \mathbb{R}^N \times \mathcal{V};\; F_\lambda(v) = 0 \big\}$$

is a C^k manifold and the natural projection $\pi : \mathcal{Z} \to \mathbb{R}^N$, $(\lambda, v) \to \lambda$ is a C^k map. Since $\dim \mathcal{Z} - N = \dim \boldsymbol{V} - \dim \boldsymbol{U}$ we deduce from Sard's theorem that most $\lambda \in \mathbb{R}^N$ are regular values of π. One can show that for such λ, 0 is a regular value of F_λ. Thus, a regularity assumption together with an ampleness condition on the family guarantee that 0 is generically a regular value of F_λ.

However, we cannot expect such genericity assuming only C^1-regularity.[2] For example, H. Whitney [158] has constructed a C^1-function $f : \mathbb{R}^2 \to \mathbb{R}$ whose set of critical values contains a nontrivial interval centered at 0.

Consider the random Gaussian function $X + f$, X standard normal random variable, Then the probability that 0 is a regular value of $X + f$ is < 1.

The above geometric argument has a probabilistic counterpart. Fix independent standard normal random variables $\Lambda_1, \dots, \Lambda_N$ and form the random Gaussian map

$$F_\omega = \sum_k \Lambda_k(\omega) F_k,$$

where $F_k \in C^r\big(\mathscr{V}, \boldsymbol{U}\big)$, $r > \dim \boldsymbol{V} - \dim \boldsymbol{U}$.

Equivalently, consider the standard Gaussian measure on $\Lambda = \mathbb{R}^N$ and think of Λ as a probability space and of F as a random map

$$F : \Lambda \times \mathscr{V} \to \boldsymbol{U}, \;\; (\lambda, \boldsymbol{v}) = F_\lambda(\boldsymbol{v}) = \sum_k \lambda_k F_k(\boldsymbol{v}).$$

Observe that a sufficient condition for F to satisfy the ampleness condition ($*$) is that for any $v \in \mathscr{V}$ we have

$$\boldsymbol{U} = \operatorname{span}\big\{ F_1(v), \dots, F_N(v) \big\}.$$

This condition also implies that F, viewed as a Gaussian random map, is ample.

We deduce from Sard's theorem that for Lebesgue almost every $\lambda \in \Lambda$, the point $0 \in \boldsymbol{U}$ is a regular value of

$$F_\lambda = \sum_{k=1}^N \lambda_k F_k.$$

[2] I am indebted to Michele Stecconi for pointing out this fact.

This implies that 0 a regular value of F_λ for λ in a set of Gaussian probability 1.

This argument was recently generalized by A. Lerario and M. Stecconi [89] as follows. Denote by E the Fréchet space $E = C^r\big(\mathscr{V}, \boldsymbol{U}\big)$. Fix a Gaussian measure Γ on E, denote by H_Γ the associated Cameron-Martin space and by S_Γ its closure in E. Equivalently, S_Γ is the topological support of Γ. In particular, $\Gamma\big[\,S_\Gamma\,\big] = 1$. Assuming that $0 \in \boldsymbol{U}$ is a regular value of the map

$$\mathbf{Ev} : S_\Gamma \times \mathcal{V} \to \boldsymbol{U},\ \ \mathbf{Ev}(F, v) = F(v),$$

then

$$\Gamma\big[\,\{0 \ \text{ is a a regular value of } F\}\,\big] = 1.$$

In Theorem 2.1.3 we approached generic regularity using a different approach. Let N be a (large) positive integer and suppose that, for each $v \in \mathcal{V}$ the collection of C^1-maps

$$\big\{\,F_k(v), F'(v)_k^\top\,\big\}_{1\le k\le N}$$

spans the vector space $\boldsymbol{U} \times \operatorname{Hom}(\boldsymbol{U}, \boldsymbol{V})$. If we define

$$F_\lambda := \sum_{k=1}^N \lambda_k F_k,\ \ \lambda = (\lambda_1, \dots, \lambda_N) \in \mathbb{R}^N,$$

then we see that the family $(F_\lambda)_{\lambda\in\mathbb{R}^N}$ satisfies ($*$). However, if $\dim \boldsymbol{V} - \dim \boldsymbol{U} > 1$, then the maps F_λ have less regularity than required by Sard's theorem.

□

2.2. The Gaussian Kac-Rice formula

We now have all the ingredients needed to prove the Gaussian Kac-Rice formula. We start by stating and proving several local versions of this version and then we will explain how these local results can be patched together to obtain a global version.

2.2.1. Local Kac-Rice formula. Suppose that $\boldsymbol{U}, \boldsymbol{V}$ are real Euclidean spaces of the same dimension m and $\mathscr{V}$ is an open subset of $\boldsymbol{V}$. We will investigate the zero sets of Gaussian C^1-maps $F : \mathscr{V} \to \boldsymbol{U}$. Before we do this we need to describe some basic properties of such zero sets of *deterministic* C^1-maps. We need to introduce a bit of terminology.

A compact subset $B \subset \boldsymbol{V}$ is called a *box subordinated to the Euclidean coordinates* $(v^1, \dots, v^m)$ on $\boldsymbol{V}$ if there exist real numbers $a_k \le b_k$, $k = 1, \dots m$, such that

$$B = \big\{\,v;\ \ v^k \in [a_k, b_k],\ \ \forall k = 1, \dots m\,\big\}.$$

It is called *nondegenerate* if $a_k < b_k$, $\forall k$. A subset is called a *box* if it is a box subordinated to a choice of Euclidean coordinates.

For any map $F : \mathcal{V} \to \boldsymbol{U}$ and any Borel set $S \subset \mathscr{V}$

$$Z[S, F] := \#\big\{\,v \in S;\ \ F(v) = 0\,\big\}.$$

Lemma 2.2.1 (Continuity of roots). *Suppose $B \subset \mathscr{V}$ is a box, and $F : \mathcal{V} \to \boldsymbol{U}$ is a C^1-map satisfying the following conditions.*

(i) $0 \in \boldsymbol{U}$ *is a regular value of F.*

(ii)
$$r_0 := \inf_{v\in\partial B} \|f(v)\| > 0.$$

Suppose that $(F_\nu)_{\nu\in\mathbb{N}}$ *is a sequence of* C^1*-maps that converge in* $C^1(\mathcal{V},\boldsymbol{U})$ *to* F*. Then*

$$\lim_{\nu\to\infty} Z[B,F_\nu] = Z[B,F] < \infty.$$

Proof. Since 0 is a regular value and $F^{-1}(0)\cap\partial B=\emptyset$ we deduce from the inverse function theorem function that F has a finite number of zeros in B, none of them located on ∂B. Set

$$\mathcal{Z} = F^{-1}(0)\cap B = \{v_1,\dots,v_n\},\ \ n=\#\mathcal{Z},\ \ \mathcal{Z}_\nu = F_\nu^{-1}(0)\cap B.$$

Using the inverse function theorem we can choose $\delta>0$ sufficiently small such that

- The open ball $B_\delta(v_i)$ is contained in B, $\forall i$,
- the closures of the balls $B_\delta(v_i)$ are disjoint, and
- the restriction of F to each of the open balls $B_\delta(v_i)$ is a diffeomorphism onto its image.

Set

$$C := B\setminus\bigcup_{i=1}^n B_\delta(v_i),\ \ r_0 := \inf_{v\in C}\|F(v)\|.$$

Since F_ν converges uniformly to F on the compact set C we deduce that there exists $\nu_0>0$ such that

$$\forall\nu\geq\nu_0,\ \ \inf_{v\in C}\|F_\nu(v)\| > r_0/2>0.$$

Thus, for $\nu\geq\nu_0$

$$\mathcal{Z}_\nu\subset\bigcup_{i=1}^n B_\delta(v_i).$$

Set

$$\mathcal{Z}_{\nu,i} := \mathcal{Z}_\nu\cap B_\delta(v_i).$$

We claim that for each $i=1,\dots,n$, there exists $\nu_i>0$ such that $\#\mathcal{Z}_{\nu,i}=1$, $\forall\nu\geq\nu_i$. We argue by contradiction. Suppose that there exists a subsequence $\mathcal{Z}_{\nu_k,i}$ such that $\#\mathcal{Z}_{\nu_k,i}\geq 2$. To ease the notation we will write $\mathcal{Z}_{k,i}$ instead of $\mathcal{Z}_{\nu_k,i}$.

Let $v_{0,k},v_{1,k}\in\mathcal{Z}_{k,i}$, $v_{0,k}\neq v_{1,k}$. Upon extracting subsequences we can assume that $v_{0,k}$ and $v_{1,k}$ converge to $v_{0,\infty},v_{1,\infty}\in \boldsymbol{cl}\,B_\delta(v_i)$. Clearly $F(v_{0,\infty})=F(v_{1,\infty})=0$ and, since F has a single zero, v_i, in $\boldsymbol{cl}\,B_\delta(v_i)$ we deduce

$$v_{0,k},\ v_{1,k}\to v_i\ \text{ as } k\to\infty.$$

Consider the unit vectors

$$w_k := \frac{1}{\|v_{1,k}-v_{0,k}\|}\big(\,v_{1,k}-v_{0,k}\,\big).$$

Upon extracting a subsequence we can assume that w_k converges to the unit vector w. Since the differential $F'(v_i)$ is invertible we deduce that $F'(v_i)w\neq 0$. Choose a linear functional $\xi:\boldsymbol{U}\to\mathbb{R}$ such that

$$\xi\big(\,F'(v_i)w\,\big)=1. \tag{2.2.1}$$

Consider now the scalar functions $f_k(v):=\xi\big(\,F_{\nu_k}(v)\,\big)$. From the mean value theorem we deduce that there exists a point p_k on the line segment $[v_{0,k},v_{1,k}]$ such that

$$0=f_k(v_{1,k})-f_k(v_{0,k})=\|v_{1,k}-v_{0,k}\|df_k(p_k)\big(\,w_k\,\big)=\|v_{1,k}-v_{0,k}\|\xi\big(\,F'_{\nu_k}(p_k)w_k\,\big).$$

In other words

$$\xi\big(\,F'_{\nu_k}(p_k)w_k\,\big)=0,\;\;\forall k.$$

Note that $p_k\to v_i$. Letting $k\to\infty$ we deduce $\xi\big(\,F'(v_i)w\,\big)=0$. This contradicts (2.2.1). □

Corollary 2.2.2 (Kac's counting formula)**.** *Suppose that $F:\mathcal{V}\to\boldsymbol{U}$ is a C^1-map and the box B satisfy the assumptions in Lemma 2.2.1. For $v\in\mathcal{V}$ we denote by $F'(v)$ the differential of F at v and by $J_F(v)$ its Jacobian $J_F=\det\big(\,F'(v)F'(v)^*\,\big)$. For $r>0$ we set*

$$Z[B,F,r]:=\frac{1}{\boldsymbol{\omega}_d r^d}\int_B \boldsymbol{I}_{\{|F|<r\}}J_F(v)dv.$$

Then, for $r>0$ sufficiently small we have $Z[B,F]=Z[B,F,r]$. In other words

$$Z[B,F]=\lim_{r\searrow 0}Z[B,F,r]. \tag{2.2.2}$$

Proof. For $u\in\boldsymbol{U}$ we set $F_u:=F-u$ so that $F_0=F$. Using Lemma 2.2.1 we deduce that

$$\lim_{u\to 0}Z[B,F_u]=Z[B,F].$$

There exists $r_0>0$ such that

$$\#F^{-1}(u)\cap B=\#F^{-1}(0)\cap B,\;\;\forall\|u\|<r_0.$$

Using the coarea formula (A.1.14) we deduce that

$$\int_{B\cap\{\|F\|<r\}}J_F(v)dv=\int_{\{\|u\|<r\}}\#F^{-1}(u)du=\boldsymbol{\omega}_d r^d\times\#F^{-1}(0)\cap B.$$

□

Corollary 2.2.3. *Suppose that $X:\Omega\times\mathscr{V}\to\boldsymbol{U}$ is an ample C^1 Gaussian field, i.e., the Gaussian vector $X(v)$ is nondegenerate for any v. Then for any box $B\subset\mathscr{V}$ the map $\omega\mapsto Z[B,X_\omega]$ is measurable.*

Proof. The map $\omega\mapsto(X_\omega,X'_\omega)$ is measurable. Since 0 is a.s. a regular value of X we deduce from Kac's counting formula that

$$Z[B,X,1/n]\to Z[B,X]\;\;\text{a.s.}.$$

Since X satisfies the standard conventions the function $Z[B,X,1/n]$ is measurable so $Z[B,X]$ is measurable as a.s. limit of measurable functions defined on a complete probability space. □

Corollary 2.2.4. *Fix a box $B\subset\mathcal{V}$. Suppose that $X_n:\Omega\times\mathcal{V}\to\boldsymbol{U}$ is a sequence of Gaussian C^1-random fields such that $X_n(v)$ is a nondegenerate Gaussian vector for any n and any $v\in\mathscr{V}$ and $X_n\to X$ a.s. in $C^1\big(\,\mathscr{V},\boldsymbol{U}\,\big)$. Then*

$$Z[B,X_n]\to Z[B,X]\;\;\text{a.s.}.$$

□

✍ *For any Borel subset S of an Euclidean space and any compactly supported continuous function $\varphi:\boldsymbol{V}\to\mathbb{R}$ we set*

$$\int_S\varphi(v)dv:=\int_S\varphi(v)\boldsymbol{\lambda}\big[\,dv\,\big],$$

where $\boldsymbol{\lambda}$ is the Lebesgue measure. This apparent excess of pedantism is fully justified. Soon we will replace $\boldsymbol{V}$ with a manifold and the measure $\boldsymbol{\lambda}$ will have to be replaced with the measure determined by a 1-density on the manifold. The above convention is meant to keep the reader alert.

Theorem 2.2.5 (Local Kac-Rice formula)**.** *Let $\boldsymbol{U}$ and $\boldsymbol{V}$ be Euclidean spaces of the same dimension m and $\mathscr{V} \subset \boldsymbol{V}$ open. Suppose that*

$$X : \Omega \times \mathcal{V} \to \boldsymbol{U},\ \ \Omega \times \mathcal{V} \ni (\omega, v) \mapsto X_\omega(v) \in \boldsymbol{U}$$

is a Gaussian C^1-field satisfying the ampleness condition (A_0), i.e., $X(v)$ is a nondegenerate Gaussian vector for any $v \in \mathscr{V}$.

If $B \subset \mathcal{V}$ is a box, then

$$\mathbb{E}\big[\, Z[B,X]\,\big] = \int_B \mathbb{E}\big[\, J_X(v) \big|\, X(v) = 0\,\big] p_{X(v)}(0) dv < \infty, \tag{KR}$$

where $J_X(v)$ denotes the Jacobian of X at $v \in \mathcal{V}$ and $\mathbb{E}\big[\, J_X(v) \big|\, X(v) = 0\big]$ denotes the conditional expectation of $J_X(v)$ given that $X(v) = 0$. We will refer to the function

$$v \mapsto \rho_{KR}(v) = \mathbb{E}\big[\, J_X(v) \big|\, X(v) = 0\,\big] p_{X(v)}(0)$$

as the Kac-Rice density *of X.*

Proof. We follow the approach in [**12**, Sec. 6.1] and [**7**, Sec. 5]. We will need the following technical result.

Lemma 2.2.6. *Denote by $\mathfrak{X}$ the space $C^1(\mathcal{V}, \boldsymbol{U})$ equipped with the topology of uniform convergence on compacts of maps and their first order derivatives.*

For any $u_0 \in \boldsymbol{U}$, $v_0 \in \mathcal{V}$, and any bounded continuous function $\alpha : \mathfrak{X} \to \mathbb{R}$ the conditional distribution $\mathbb{P}_{\alpha(X)|X(v_0)=u_0}$ is well defined as a probability measure on $\mathbb{R}$ and depends continuously on u_0 in the topology of weak convergence of measures.

Proof. From Proposition 1.1.32 (Gaussian regression formula)we we deduce that any $v \in V$ we have

$$X(v) = R_{X(v),X(v_0)} X(v_0) + Z(v, v_0),$$

where the random variable $Z(v, v_0)$ is independent of $X(v_0)$ and the regression operator $R_{X(v),X(v_0)}$ is given by (1.1.16). We have

$$Z_\omega(v, v_0) = X_\omega(v) - R_{X(v),X(v_0)} X_\omega(v_0).$$

Hence, for any ω the map $v \mapsto Z_\omega(v, v_0)$ is also C^1. The resulting map $\mathcal{V} \ni v_0 \mapsto Z_\omega(-, v_0) \in \mathfrak{X}$ is continuous for any ω.

Fix a continuous and bounded function $\alpha : \mathfrak{X} \to \mathbb{R}$. Then the real number

$$\alpha(X_\omega) = \alpha\big(\, R_{X_\omega(-),X_\omega(v_0)} u_0 + Z_\omega(-, v_0)\,\big)$$

depends continuously on (u_0, v_0) for any ω and, since α is bounded, we deduce from the Dominated Convergence Theorem and the regression formula that

$$\mathbb{E}\big[\, \alpha(X) \big|\, X(v_0) = u_0\,\big] = \mathbb{E}\big[\, \alpha\big(\, R_{X,X(v_0)} u_0 + Z(-, v_0)\,\big)\,\big]$$

depends continuously on (u_0, v_0). □

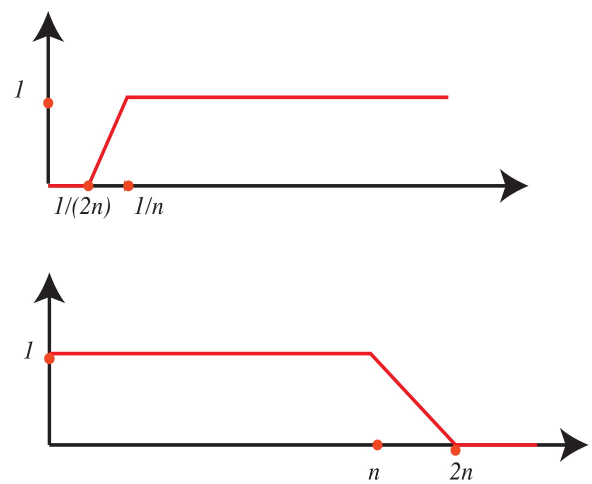

Figure 2.1. *The graphs of F_n (top) and the graph of G_n (bottom).*

Let $F : [0,\infty) \to [0,1]$ be the continuous piecewise linear function such that

$$F(x) = \begin{cases} 0, & x \leq 1/2, \\ 1, & x \geq 1. \end{cases}$$

For $n \in \mathbb{N}$ we set

$$F_n(x) := F(nx), \quad G_n(x) = 1 - F(x/(2n)). \tag{2.2.3}$$

The functions F_n and G_n are depicted in Figure 2.1. For $v \in \mathcal{V}$ we set

$$d_v := \operatorname{dist}\big(\, v, \partial B\,\big),$$

and we denote by J_v the Jacobian of X at v. For $u \in \boldsymbol{U}$ and $n \in \mathbb{N}$ and $\Phi \in C^1(\mathscr{V}, \boldsymbol{U})$ we set

$$C_u^n(\Phi, B) := \sum_{v \in \Phi^{-1}(u) \cap B} F_n\big(\, J_\Phi(v)\,\big) G_n\big(\, J_\Phi(v)\,\big) F_n\big(\, d_v\,\big).$$

Lemma 2.2.7. *Let $\mathfrak{X} = C^1(\mathscr{V}, \boldsymbol{U})$. Then the following hold.*

(i) *For any $u \in \boldsymbol{U}$ the map $\mathfrak{X} \ni \Phi \mapsto C_u^n(\Phi, B)$ is continuous*

(ii) *For any $\Phi \in \mathfrak{X}$ the map $u \mapsto C_u^n(\Phi, B)$ is continuous.*

□

We proceed assuming the validity of the above lemma. We set

$$C_u^n(B) := C_u^n(X, B) = \sum_{v \in X^{-1}(u) \cap B} F_n\big(\, J_v\,\big) G_n\big(\, J_v\,\big) F_n\big(\, d_v\,\big), \tag{2.2.4}$$

$$Q_u^n(B) := C_u^n(B) G_n\big(\, C_u^n(B)\,\big).$$

These are measurable as compositions of measurable functions $\mathfrak{X} \to \mathbb{R}$ with $X : \Omega \to \mathfrak{X}$.

Note that $C_u^n(B)$ is the number of solutions v of the equation $X(v) = u$ in the compact (random) set

$$K_n := \Big\{ v \in B : \;\; J_v, d_v \geq \frac{1}{2n}, \;\; J_v \leq 2n \Big\}.$$

Intuitively, $C_u^n(B)$ counts the solutions v of $X(v) = u$ located in B for which that Jacobian J_v is not too small, not too large and they are not too close to the boundary of B. The quantity $Q_u^n(B)$ is a sort of truncation of $C_u^n(B)$. Note that $Q_u^n(B) = 0$ whenever $C_u^n(B) > n$.

Let $g : \boldsymbol{U} \to [0, \infty)$ be a continuous, compactly supported function. The coarea formula (A.1.14) implies that

$$\int_{\boldsymbol{U}} g(u) Q_u^n(B) du = \int_B J_v F_n(J_v) G_n(J_v) F_n(d_v) G_n\big(C_{X(v)}^n(B) \big) g\big(X(v) \big) dv.$$

The standard assumptions guarantees that the above quantities are measurable. These random variables are bounded since the various integrands are bounded. E.g., $Q_u^n(B) \leq 2n$. Taking expectations we deduce

$$\int_{\boldsymbol{U}} g(u) \mathbb{E}\big[Q_u^n(B) \big] du = \int_B \mathbb{E}\big[J_v F_n(J_v) G_n(J_v) F_n(d_v) G_n\big(C_{X(v)}^n(B) \big) g\big(X(v) \big) \big] dv$$

$$= \int_{\boldsymbol{U}} g(u) \left(\int_B \mathbb{E}\big[J_v F_n(J_v) G_n(J_v) F_n(d_v) G_n\big(C_{X(v)}^n(B) \big) \big| \, X(v) = u \big] dv \right) p_{X(v)}(u) du.$$

Since the above equality holds for any continuous compactly supported function g we deduce

$$\mathbb{E}\big[Q_u^n(B) \big] = \int_B \mathbb{E}\big[J_v F_n(J_v) G_n(J_v) F_n(d_v) G_n\big(C_{X(v)}^n(B) \big) \big| \, X(v) = u \big] p_{X(v)}(u) dv \tag{2.2.5}$$

for *almost every* $u \in \boldsymbol{U}$. To prove that the above equality holds *for any* u we will show that both sides of (2.2.5) depend continuously on u.

The random function $u \mapsto C_u^n(B) = C_u^n(X, B)$ is a.s. continuous since

$$u \mapsto C_u^n(\Phi)$$

is continuous for any $\Phi \in \mathfrak{X}$. Consider

$$\alpha_v^n : \mathfrak{X} \to \mathbb{R}, \;\; \alpha_v(\Phi) := \underbrace{J_\Phi(v) F_n(J_\Phi(v)) G_n(J_\Phi(v))}_{\leq 2n} F_n(d_v) G_n\big(C_{\Phi(v)}^n \big).$$

For fixed v it depends continuously with respect to Φ in the topology of $\mathfrak{X}$. We can rewrite the right-hand-side of (2.2.5) as

$$\int_B \mathbb{E}\big[\alpha_v^n(X) \big| \, X(v) = u \big] p_{X(v)}(u) dv.$$

Corollary 2.2.4 and Lemma 2.2.6 show that the integrand depends continuously on u. Clearly it is bounded uniformly in u. The Dominated Convergence Theorem shows that the above integral depends continuously on u. Hence

$$\mathbb{E}\big[Q_u^n(B) \big] = \int_B \mathbb{E}\big[J_v F_n(J_v) G_n(J_v) F_n(d_v) G_n\big(C_{X(v)}^n(B) \big) \big| \, X(v) = u \big] p_{X(v)}(u) dv, \tag{2.2.6}$$

for *every* $u \in \boldsymbol{U}$. In particular, for $u = 0$ we deduce

$$\mathbb{E}\big[Q_0^n(B) \big] = \int_B \mathbb{E}\big[J_v F_n(J_v) G_n(J_v) F_n(d_v) G_n\big(C_{X(v)}^n(B) \big) \big| \, X(v) = 0 \big] p_{X(v)}(0) dv. \tag{2.2.7}$$

Bulinskaya's Lemma 2.1.2 and the Transversality Theorem 2.1.4 imply that 0 is a.s. a regular value of X and the equation $X(v) = 0$ has no solutions on ∂B. We deduce that

$$Q_0^n(B) \nearrow Z[B, X] \text{ as } n \to \infty.$$

Since $F_n, G_n \nearrow 1$ we can use the Monotone Convergence Theorem in (2.2.7) as $n \to \infty$ and deduce (**KR**) assuming the validity of Lemma 2.2.7. Observe that Lemma 2.2.6 shows that the map

$$B \ni v \mapsto \mathbb{E}\big[\, J_X(v) \big|\, X(v) = 0\,\big] \in \mathbb{R}$$

is continuous and, since $X(v)$ is nondegenerate for any v, we deduce that

$$\int_B \mathbb{E}\big[\, J_X(v) \big|\, X(v) = 0\,\big] p_{X(v)}(0) dv < \infty.$$

Proof of Lemma 2.2.7. The proof is similar to the proof of Lemma 2.2.1. Fix $\Phi_0 \in \mathfrak{X}$ and $u_0 \in \boldsymbol{U}$. For each $n \in \mathbb{N}$ we consider the compact set

$$K_n := \left\{ v \in B;\ \operatorname{dist}(v, \partial B) \geq 1/n,\ \ \frac{1}{2n} \leq J_{\Phi_0}(v) \leq 2n \right\}.$$

Note that $K_n \subset \boldsymbol{int}(K_{n+1})$, $\forall n$. Let

$$\mathcal{Z}_n(\Phi_0) = \Phi_0^{-1}(u_0) \cap K_n.$$

Observe that if $v \in \mathcal{Z}_n(\Phi_0)$, then the differential $\Phi_0'(v)$ is invertible so and the inverse function theorem implies that there exists an open neighborhood $\mathcal{O}_v$ of such that $\Phi_0^{-1}(u_0) \cap \mathcal{O}_v = \{v\}$. Hence $\mathcal{Z}_n(\Phi_0)$ is a closed subset of a compact set consisting of isolated points so $Z_n(\Phi_0)$ is finite

$$\mathcal{Z}_n(\Phi_0) := \{v_1, \dots, v_n\}$$

Invoking the inverse function theorem we deduce that there exist $r > 0$ and pairwise disjoint open sets $\mathcal{O}_1, \dots, \mathcal{O}_n$ with the following properties.

- $v_k \in \mathcal{O}_k \subset \boldsymbol{int}\, K_{n+1}$, $\forall k = 1, \dots, n$. We set
$$\mathcal{O} := \bigcup_{k=1}^n \mathcal{O}_k.$$
- The restriction of Φ_0 to $\mathcal{O}_k$ is a diffeomorphism onto the open ball $B_r(u_0) \subset \boldsymbol{U}$.

Suppose that $\|\Phi_\nu - \Phi_0\|_{C^1(B)} \to 0$ as $\nu \to \infty$. We claim that

$$\exists N > 0 :\ \ \forall \nu \geq N,\ \ \Phi_\nu^{-1}(u_0) \cap K_n \subset \mathcal{O}.$$

We argue by contradiction. Suppose that there exists a subsequence $\nu_m \nearrow \infty$ and and

$$w_{\nu_m} \in \Phi_{\nu_m}^{-1}(u_0) \cap K_n \setminus \mathcal{O},\ \ \forall m \tag{2.2.8}$$

Upon extracting a subsequence we can assume that w_{ν_m} converges to $w_* \in K_n$. Letting $m \to \infty$ in the equality $\Phi_{\nu_m}(w_{\nu_m}) = u_0$ we deduce $\Phi_0(w_*) = u_0 \in \mathcal{O}$. This contradicts (2.2.8).

Arguing as in the proof of Lemma 2.2.1 we conclude that there exists $N > 0$ such that for any $\nu \geq N$ and any $k = 1, \dots, \nu$ the equation $\Phi_\nu(v) = u_0$ has at most one solution $v \in \mathcal{O}_k$.

Let us now observe that for ν sufficiently large the equation $\Phi_\nu(v) = u_0$ has one solution $v \in \mathcal{O}_k$. This is an immediate consequence of the theory of degree of a continuous map; see e.g. [**122**, Chap.1]. Indeed, if $B_{r_k}(v_k)$ is a small closed ball centered at v_k and contained in $\mathcal{O}_k$, then for ν sufficiently large

$$\sup_{v\in\partial B_{r_k}(v_k)} \|\Phi_\nu(v)-u_0\|>0$$

and

$$\pm 1=\deg\big(\Phi, B_{r_k}(v_k),0\big)=\lim_{\nu\to\infty}\deg\big(\Phi_\nu, B_{r_k}(v_k),0\big).$$

This proves that for any continuous function $\varphi: B\to\mathbb{R}$ such that $\operatorname{supp}\varphi\subset K_n$ we have

$$\lim_{\nu\to\infty}\sum_{v\in\Phi_\nu^{-1}(u_0)}\varphi(v)=\sum_{v\in\Phi_0^{-1}(u_0)}\varphi(v).$$

This proves the first part of Lemma 2.2.7. The second follows from the above first part applied to the maps $\Phi_\nu=\Phi_0-(u_\nu-u_0)$, where $u_\nu\to u_0$. □

This completes the proof of the local Kac-Rice formula □

Recall the random variable

$$Z[B,X,r]:=\frac{1}{\boldsymbol{\omega}_d r^d}\int_B \boldsymbol{I}_{\{|X|<r\}}J_X(v)dv$$

that appears in Kac's counting formula (2.2.2)

$$Z[B,X]=\lim_{r\searrow 0}Z[B,X,r].$$

Proposition 2.2.8. *Let X as in Theorem 2.2.5. Assume additionally that X is 0-ample, i.e., for any $v\in\mathcal{V}$ the Gaussian vector $X(v)$ is nondegenerate. Then*

$$\mathbb{E}\big[\,Z[B,X]\,\big]=\lim_{r\searrow 0}\mathbb{E}\big[\,Z[B,X,r]\,\big]$$

In particular $Z[B,X,r]\to Z[B,X]$ in L^1 as $r\searrow 0$.

Proof. Using Fubini's formula we deduce

$$\mathbb{E}\big[\,Z[B,X,r]\,\big]=\frac{1}{\boldsymbol{\omega}_d r^d}\int_B\mathbb{E}\big[\,\boldsymbol{I}_{\{|X|<r\}}J_X(v)\,\big]dv$$

Note that

$$\mathbb{E}\big[\,\boldsymbol{I}_{\{|X|<r\}}J_X(v)\,\big]=\int_{|u|<r}\mathbb{E}\big[\,J_X(v)\big|\,X(v)=u\,\big]p_{X(v)}(u)du,$$

so that

$$\mathbb{E}\big[\,Z[B,X,r]\,\big]=\frac{1}{\boldsymbol{\omega}_d r^d}\int_B\Big(\int_{|u|<r}\mathbb{E}\big[\,J_X(v)\big|\,X(v)=u\,\big]p_{X(v)}(u)du\Big)dv$$

$$=\frac{1}{\boldsymbol{\omega}_d r^d}\int_{|u|<r}\underbrace{\Big(\int_B\mathbb{E}\big[\,J_X(v)\big|\,X(v)=u\,\big]p_{X(v)}(u)dv\Big)}_{=:\varphi(u)}du.$$

The regression formula shows that the integrand $u\mapsto\varphi(u)$ is continuous on $|u|\le r$ so that

$$\lim_{r\searrow 0}\frac{1}{\boldsymbol{\omega}_d r^d}\int_{|u|<r}\varphi(u)du=\varphi(0)=\int_B\mathbb{E}\big[\,J_X(v)\big|\,X(v)=0\,\big]p_{X(v)}(0)dv=\mathbb{E}\big[\,Z[B,X]\,\big].$$

The statement about the L^1-convergence follows from the Lebesgue-Vitali theorem on uniform integrability; see e.g. [118, Sec. 3.2.2] or [141, Thm. 16.6]. □

Remark 2.2.9. (a) The assumptions in Theorem 2.2.5 are less stringent than in the Gaussian Kac-Rice formula [1, Cor.11.2.2]. Theorem 2.2.5 requires only the 0-ampleness of X whereas [1, Cor.11.2.2] requires J_1-ampleness. This is due to the different strategy used in [1] to prove (**KR**).

(b) There are many generalizations of the local Kac-Rice formula (**KR**). First, one can slightly relax the Gaussian assumption; see [1, 7, 147]. These generalizations do not seem too practical for two reasons. First, the various conditions imposed on the random field are difficult do verify in the non-Gaussian case. Then, the computation of the conditional expectation in the KR-density is nearly impossible in the non-Gaussian case.

There exist other generalizations, of a more geometric nature. Consider a Gaussian field

$$F : \mathscr{V} \to \boldsymbol{U}, \;\; \dim \boldsymbol{V} \leq \dim \boldsymbol{U},$$

and $M \subset \boldsymbol{U}$, a submanifold of dimension $\dim \boldsymbol{U} - \dim \boldsymbol{V}$, then one expects that the map F is a.s. transversal to M so one expects that $F^{-1}(M)$ is discrete and we can ask what is the expectation of its cardinality. This problem is discused in great detail in a much greater generality in M. Stecconi's dissertation [147].

If $m_0 := \dim \boldsymbol{U} < m_1 := \dim \boldsymbol{V}$, then the preimage $F^{-1}(0)$ is typically a submanifold of $\mathscr{V}$ codimension m_0. We can then ask what is the $\mathbb{E}\big[\, \mathcal{H}_{m_1 - m_0}(F^{-1}(0))\,\big]$, the expectation of the volume of $F^{-1}(0)$, where $\mathcal{H}_k$ denotes the k-dimensional Hausdorff measure. This situation is addressed in [7, 12]. □

2.2.2. A weighted local Kac-Rice formula. Let X be as in Theorem 2.2.5. Using [80, Lemma 3.1] and Theorem 2.2.5 we deduce that the correspondence

$$\mathcal{B}_{\mathscr{V}} \mapsto Z[S, X]$$

is a locally finite random measure, i.e.,

(i) for any $S \in \mathcal{B}_{\mathscr{V}}$ the map

$$(\Omega, \mathcal{S}, \mathbb{P}) \ni \omega \mapsto Z[S, X_\omega] \in [0, \infty]$$

is measurable, and

(ii) for any $\omega \in \Omega$ the map

$$\mathcal{B}_{\mathscr{V}} \ni S \mapsto\mapsto Z[S, X_\omega] \in [0, \infty]$$

is a measure that is finite on the compact subsets of $\mathscr{V}$.

We refer to Appendix C.2 for more information about locally finite random measures. The integral of continuous, compactly supported function $\varphi \in C^0_{\mathrm{cpt}}(\mathscr{V})$ with respect to this random measure produces a random variable

For any compactly supported measurable function $\varphi : \mathscr{V} \to \mathbb{R}$ we set

$$Z[\varphi, X] = \int_{\mathscr{V}} \varphi(v) Z[dv, X] = \sum_{v \in X^{-1}(0)} \varphi(v).$$

More generally, for any nonnegative measurable function $f : \mathscr{V} \to [0,\infty)$ we can define $Z[f,X]$ in a similar way so $Z[B,X] = Z[\boldsymbol{I}_B, X]$.

We have the following variant of the Kac-Rice formula.

Theorem 2.2.10 (Weighted local Kac-Rice formula)**.** *Let X be as in Theorem 2.2.5 and $\varphi \in C^0_{\mathrm{cpt}}(\mathscr{V})$. Then*

$$\mathbb{E}\big[\, Z[\varphi, X]\,\big] = \int_{\mathscr{V}} \varphi(v)\rho^X_{KR}(v)dv < \infty, \tag{2.2.9}$$

where

$$\rho^X_{KR}(v) = \mathbb{E}\big[\, J_X(v)\,\big|\, X(v) = 0\,\big] p_{X(v)}(0)$$

is the Kac-Rice density *of X.*

Proof. Decomposing $\varphi = \varphi_+ - \varphi_-$ we can assume that $\varphi \geq 0$. Using partitions of unity we can reduce everything to the special case when $\operatorname{supp}\varphi$ is contained in a box $B \subset \mathscr{V}$. Now run the argument in the proof of the local Kac-Rice formula with the term $C^n_u(B)$ in (2.2.4) replaced by

$$C^n_u(\varphi) := \sum_{v \in X^{-1}(u)} \varphi(v) F_n\big(\, J_v\,\big) G_n\big(\, J_v\,\big) F_n\big(\, \operatorname{dist}(v, \partial B)\,\big).$$

Note that formally $C^n_u(B) = C^n_u\big(\, \boldsymbol{I}_B\,\big)$. □

Suppose that $F : \mathscr{V} \to \mathbb{R}$ is a Morse function. In particular, F is at least C^2. We denote by dF its differential. We view it as a map

$$dF : \mathscr{V} \to \boldsymbol{V}^*.$$

We set

$$\mathfrak{C}[-, F] := Z[-, dF].$$

Hence, for any Borel subset $S \subset \mathscr{V}$,

$$\mathfrak{C}\big[\, S, F\,\big] = \#\big\{\, v \in S;\ \ dF(v) = 0\,\big\}.$$

For any continuous function $\varphi : B \to \mathbb{R}$ we set

$$\mathfrak{C}\big[\, \varphi, F\,\big] = \int_{\mathscr{V}} \varphi(v)\mathfrak{C}[dv, F] = \sum_{dF(v)=0} \varphi(v),$$

Corollary 2.2.11. *Suppose that $F : \mathscr{V} \to \mathbb{R}$ is a C^2 Gaussian function satisfying the* 1*-ampleness condition*

> *for any $v \in \mathscr{V}$ the Gaussian vector*
>
> $$\Omega \ni \omega \mapsto dF_\omega \in \boldsymbol{V}^*$$
>
> *is nondegenerate.*

Then F is a.s. *Morse and for any function $\varphi \in C^0_{\mathrm{cpt}}(\mathscr{V})$ we have*

$$\begin{aligned} &\mathbb{E}\big[\, \mathfrak{C}[\varphi, F]\,\big] = \mathbb{E}\big[\, Z[\varphi, dF]\,\big] \\ &= \int_{\mathscr{V}} \mathbb{E}\big[\, |\det \operatorname{Hess}_F(v)|\,\big|\, dF(v) = 0\,\big] p_{dF(v)}(0)\varphi(v)dv < \infty. \end{aligned} \tag{2.2.10}$$

The quantity $\mathbb{E}\big[\, |\det \operatorname{Hess}_F(v)|\,\big|\, dF(v) = 0\,\big] p_{dF(v)}(0)dv$ is the Kac-Rice density of dF □

We conclude this subsection with another version of the local Kac-Rice formula which counts the zeros of a *random* map inside a *random* set.

Theorem 2.2.12. *Let $\boldsymbol{U}$, $\boldsymbol{V}$ and $\boldsymbol{W}$ be Euclidean spaces such that* $\dim \boldsymbol{U} = \dim \boldsymbol{V} = m$. *Let $\mathscr{V} \subset \boldsymbol{V}$ be open. Suppose that*

$$X \oplus Y : \Omega \times \mathcal{V} \to \boldsymbol{U} \oplus \boldsymbol{W}, \;\; \Omega \times \mathcal{V} \ni (\omega, v) \mapsto X_\omega(v) \oplus Y_\omega \in \boldsymbol{U} \oplus \boldsymbol{W},$$

is a C^1-Gaussian field satisfying the standard conditions such that, for any $v \in \mathscr{V}$, the Gaussian vector

$$\Omega \ni \omega \mapsto X_\omega(v) \oplus Y_\omega(v) \in \boldsymbol{U} \oplus \boldsymbol{W}$$

is nondegenerate. For each $v \in \mathscr{V}$ we denote by J_v the Jacobian of X at v,

$$J_v = \sqrt{\det\big(\, X'(v) X'(v)^* \,\big)}.$$

Suppose that $C \subset \boldsymbol{W}$ is a nondegenerate box in $\boldsymbol{W}$. Then, for any box $B \subset \mathscr{V}$

$$\mathbb{E}\big[\, Z[B \cap Y^{-1}(C), X)] \,\big] = \int_B \mathbb{E}\big[\, J_v \boldsymbol{I}_C\big(\, Y(v)\,\big) \big|\, X(v) = 0 \,\big] p_{X(v)}(0) dv \tag{2.2.11a}$$

$$= \int_B \left(\int_C \mathbb{E}\big[\, J_v \big|\, X(v) = 0, Y(v) = w \,\big] \mathbb{P}_{Y(v)}\big[\, dw \,\big] \right) p_{X(v)}(0) dv. \tag{2.2.11b}$$

Proof. We know that a.s. the equation $X(v) = 0$ has no solutions on ∂B and 0 is a regular value of X. Let us show that a.s. there exist no solution of of $X(v) = 0$ such that $Y(v) \in \partial C$.

Consider the Gaussian field

$$F : \mathscr{V} \oplus \boldsymbol{W} \to \boldsymbol{U} \oplus \boldsymbol{V}, \;\; (v, w) \mapsto \big(\, X(v), Y(v) - w \,\big).$$

Since the Gaussian vector $X(v) \oplus Y(v)$ is nondegenerate, so is $F(v, w)$. Set $K := B \times \partial C$.

The Hausdorff dimension of K is $< \dim(\boldsymbol{V} \oplus \boldsymbol{W})$ and Bulinskaya's Lemma 2.1.2 implies that a.s. the equation $F(v, w) = 0$ has no solution in K. Thus, with probability 1, there exists no $v \in B$ such that $X(v) = 0$ and $Y(v) \in \partial C$.

Denote by C° the interior of C and by E the complement of C° in $\boldsymbol{W}$. We set

$$C_n := \big\{\, w \in \boldsymbol{W};\;\; \mathrm{dist}(w, E) \geq 1/n \,\big\} \;\text{ and }\; \eta_n(w) = \frac{\mathrm{dist}(w, E)}{\mathrm{dist}(w, C_n) + \mathrm{dist}(w, E)}.$$

Note that $\eta_n(w) \nearrow \boldsymbol{I}_C(w)$, $\forall w \in \boldsymbol{W}$. Thus as $n \to \infty$

$$\sum_{v \in X^{-1}(0) \cap B} \eta_n\big(\, Y(v)\,\big) \;\nearrow\; \# X^{-1}(0) \cap B \cap Y^{-1}(C), \;\; \text{a.s.}.$$

The fact that almost surely X has no zero in $B \cap Y^{-1}(\partial C)$ plays a key role in the above equality.

For $v \in \mathscr{V}$, $u \in \boldsymbol{U}$, $n \in \mathbb{N}$, $\Phi \in C^1(\mathscr{V}, \boldsymbol{U})$ and $\Psi \in C^1(\mathscr{V}, \boldsymbol{W})$ we set

$$d_v := \mathrm{dist}(v, \partial B),$$

$$C^n_u(\Phi, \Psi, B) := \sum_{v \in \Phi^{-1}(u) \cap B} \eta_n\big(\, \Psi(v)\,\big) F_n\big(\, J_\Phi(v)\,\big) G_n\big(\, J_\Phi(v)\,\big) F_n\big(\,, d_v\,\big)$$

where F_n and G_n are defined by (2.2.3). An immediate modification of the proof of Lemma 2.2.7 yields the following continuity result.

Lemma 2.2.13. *The functions*

$$C^1(\mathscr{V},\boldsymbol{U})\times C^1(\mathscr{V},\boldsymbol{W})\ni(\Phi,\Psi)\mapsto C^n_u(\Phi,\Psi,B)$$

and $u\mapsto C^n_u(\Phi,\Psi)$ *are continuous.* □

We set

$$C^n_u(B):=C^n_u(X,Y,B)=\sum_{v\in X^{-1}(u)\cap B}\eta_n\big(Y(v)\big)F_n\big(J_v\big)G_n\big(J_v\big)F_n\big(d_v\big),\tag{2.2.12}$$

$$Q^n_u(B):=C^n_u(B)G_n\big(C^n_u(B)\big).$$

These are measurable as compositions of continuous functions $C^1(\mathscr{V},\boldsymbol{U})\times C^1(\mathscr{V},\boldsymbol{W})\to\mathbb{R}$. The argument in the proof of Theorem 2.2.5 now carries through without any conceptual changes and yields (2.2.11a). □

The argument in the above proof can be used to produce the following version of Theorem 2.2.12.

Corollary 2.2.14. *Let* $\boldsymbol{U}$, $\boldsymbol{V}$ *and* $\boldsymbol{W}$ *be Euclidean spaces such that* $\dim\boldsymbol{U}=\dim\boldsymbol{V}=m$. *Let* $\mathscr{V}\subset\boldsymbol{V}$ *be open. Suppose that*

$$X:\Omega\times\mathcal{V}\to\boldsymbol{U},\;\;\Omega\times\mathcal{V}\ni(\omega,v)\mapsto X_\omega(v)\in\boldsymbol{U},$$

$$Y:\Omega\times\mathcal{V}\to\boldsymbol{W},\;\;\Omega\times\mathcal{V}\ni(\omega,v)\mapsto X_\omega(v)\in\boldsymbol{U},$$

are C^1*-Gaussian fields such that, for any* $v\in\mathscr{V}$ *the Gaussian vector*

$$\Omega\ni\omega\mapsto X_\omega\oplus Y_\omega\in\boldsymbol{U}\oplus\boldsymbol{W}.$$

is nondegenerate. For any any functions $f\subset C^0_{\mathrm{cpt}}(\mathscr{V})$, $g\in C^0_{\mathrm{cpt}}(W)$ *we set*

$$Z[f,g;X,Y]=\sum_{v\in X^{-1}(0)}f(v)g\big(Y(v)\big).$$

Then

$$\mathbb{E}\big[\,Z[f,g;X,Y]\,\big]=\int_{\mathscr{V}}\mathbb{E}\big[\,J_vg\big(Y(v)\big)\big|\,X(v)=0\,\big]f(v)p_{X(v)}(0)dv\tag{2.2.13a}$$

$$=\int_{\mathscr{V}}\left(\int_{\boldsymbol{W}}\mathbb{E}\big[\,J_v\big|\,X(v)=0,Y(v)=w\,\big]g(w)\mathbb{P}_{Y(v)}\big[\,dw\,\big]\right)f(v)p_{X(v)}(0)dv\tag{2.2.13b}$$

□

Suppose that $F:\mathscr{V}\to\mathbb{R}$ is a Morse function and $B\subset\mathscr{V}$ is a nondegenerate box. We denote by $\mathcal{D}(F|_B)$ the *discriminant set* of $F|_B$, i.e., the set of critical values of $F\,|_B$. The *discriminant measure* of $F|_B$ is the pushforward

$$\mathfrak{D}_{B,F}=F_{\#}\mathfrak{C}[-,F]\,\big|_B=\sum_{t\in\mathbb{R}}\mathfrak{C}[F^{-1}(t)\cap B,F]\delta_t.$$

The discriminant measure is concentrated on $\mathcal{D}(F|_B)$. For $\varphi\in C^0_{\mathrm{cpt}}(\mathbb{R})$ we set

$$\mathfrak{D}_{B,F}\big[\,\varphi\,\big]:=\int_{\mathbb{R}}\varphi(t)\mathfrak{D}_{B,F}[dt]=\sum_{\substack{dF(v)=0,\\ v\in B}}\varphi\big(\,F(v)\,\big).$$

When F is random, $D_{B,F}\big[\,\varphi\,\big]$ is a random variable.

Corollary 2.2.15. *Suppose that $F:\mathscr{V}\to\mathbb{R}$ is a C^2 Gaussian function satisfying the J_1-ampleness condition*

> *for any $v\in\mathscr{V}$ the Gaussian vector*
> $$\Omega\ni\omega\mapsto F_\omega\oplus dF_\omega\in\mathbb{R}\oplus\boldsymbol{V}^*$$
> *is nondegenerate.*

Then F is a.s. *Morse and, for any box $B\subset\mathscr{V}$ and any function $\varphi\in C^0_{\mathrm{cpt}}(\mathbb{R})$ we have*

$$\begin{gathered}\mathbb{E}\big[\,\mathfrak{D}_{B,F}[\varphi]\,\big]=\\ =\int_B\left(\int_{\mathbb{R}}\mathbb{E}\big[\,\big|\det\mathrm{Hess}_F(v)\big|\,\big|\,dF(v)=0,\;F(v)=t\,\big]\varphi(t)\mathbb{P}_{F(v)}[dt]\right)p_{dF(v)}(0)dv.\end{gathered}\tag{2.2.14}$$

□

2.2.3. Global Kac-Rice formula. Suppose that (M,g) is smooth, compact[3] connected m-dimensional Riemannian manifold and $E\to M$ is a smooth real vector bundle of rank $m=\dim M$. We assume additionally that E is equipped with a metric h and a connection ∇^E compatible with the metric h. Denote by $\mathrm{Vect}(M)$ the space of smooth vector fields on M and by vol_g the volume measure on M determined by the Riemann metric g.

We set $\mathrm{End}(E):=E\otimes E^*\to M$. This is a smooth vector bundle over M whose fiber over a point $x\in M$ is

$$E_x\otimes E_x^*\cong\mathrm{End}(E_x).$$

Suppose that $u:M\to E$ is a C^1-section of E. For each point $x_0\in M$ we have a linear map

$$\nabla^E u:T_{x_0}M\to E_{x_0},\;\;T_{x_0}M\ni v\mapsto\nabla^E_v u(x_0):=\nabla^E_{X_v}u(x_0),$$

where $X_v\in\mathrm{Vect}(M)$ is any smooth vector field on M such that $X_v(x_0)=v$. Let us observe that if $u(x_0)=0$ and ∇^0,∇^1 are two connections on E, then

$$\nabla^0u(x_0)=\nabla^1u(x_0).$$

Indeed, if we set $A=\nabla^1-\nabla^0$, then A is a section of $T^*M\otimes\mathrm{End}(E)$. Then

$$\nabla^1_v u(x_0)-\nabla^0_v u(x_0)=A(X_v)u(x_0)=0.$$

Thus, at every zero x_0 of the section u, the linear map

$$T_{x_0}M\to E_{x_0},\;\;v\mapsto\nabla_v u(x_0).$$

is *independent* of the connection ∇ on E. Following the custom in algebraic geometry we will refer to this map as the *adjunction map* of u at x_0 and we will denote it by $\mathrm{adj}_u(x_0)$. A zero x_0 of u is called *nondegenerate* if the adjunction map $\mathrm{adj}(x_0)$ is invertible.

Theorem 2.2.16 (Global Kac-Rice formula)**.** *Suppose that (M,g) is smooth, compact connected m-dimensional Riemannian manifold and $E\to M$ is a smooth real vector bundle of rank $m=\dim M$. Fix a smooth metric h on E and a connection ∇^E compatible with the metric h.*

[3] The compactness is not needed but it is the only situation we will deal with in this book.

Let Ψ be a Gaussian random section of $E \to M$ that is a.s. C^1 *and satisfies the ampleness condition*

$$\forall x \in M, \text{ the } E_x\text{-valued Gaussian vector } \Psi(x) \text{ is nondegenerate.} \tag{2.2.15}$$

Then the following hold.

(i) *The zeros of Ψ are* a.s.*-nondegenerate.*

(ii) *For any continuous function $\varphi : M \to \mathbb{R}$ we set*

$$Z[\varphi, \Psi] = \sum_{\Psi(x)=0} \varphi(x).$$

Then the $Z[\varphi, \Psi]$ is measurable and

$$\mathbb{E}\big[\, Z(\varphi, \Psi) \,\big] = \int_M \mathbb{E}\big[\, J_{\nabla^E \Psi(x)} \,\big|\Psi(x) = 0 \,\big] p_{\Psi(x)(0)} \varphi(x) \operatorname{vol}_g \big[\, dx \,\big]. \tag{2.2.16}$$

Above, $p_{\Psi(x)}$ denotes the probability density of the nondegenerate Gaussian vector $\Psi(x)$ and $J_{\nabla^E \Psi(x)}$ denotes the Jacobian of the linear map $\nabla^E \Psi(x) : T_x M \to E_x$ computed in terms of the inner product g_x on $T_x M$ and h_x on E_x.

Proof. Clearly, the left-hand side of (2.2.16) is independent of the various choices: the metric g on M, the metric h on E and the connection ∇^E. We first prove that the right-hand side of this equality is also independent of these choices.

1. Independence of the connection. This is easy. Given that $\Psi(x) = 0$ we have

$$\nabla^E \psi(x) = \operatorname{adj}_\Psi(x)$$

and the right-hand side is independent of any connection on E.

2. Independence of the metric g. Suppose that g^1, g^0 are two Riemann metrics on M, then there exists a smooth endomorphism S of TM that is symmetric and positive definite with respect to the metric g^0 and such that

$$g^1(X, Y) = g^0\big(\, SX, Y \,\big)$$

Then

$$\operatorname{vol}_{g^1} \big[\, dx \,\big] = \sqrt{\det S_x} \cdot \operatorname{vol}_{g^0} \big[\, dx \,\big],$$

Denote by J^i_x the jacobian of $L_x : \nabla^E u(x) : T_x M \to E_x$ computed with respect to the inner product g^i_x on $T_x M$ and the inner product h_x on E_x. The inner products g^i determine two Lebesgue measures $\boldsymbol{\lambda}_{g^i_x}$ on $T_x M$ related by the equality

$$\boldsymbol{\lambda}_{g^1_x} \big[\, dv \,\big] = \sqrt{\det S_x} \cdot \boldsymbol{\lambda}_{g^0_x} \big[\, dv \,\big].$$

The inner product h_x determines a Lebesgue measure $\boldsymbol{\lambda}_{h_x}$ on E_x. The equality (A.1.6) shows that

$$\boldsymbol{\lambda}_{h_x} = J^i_x \cdot (L_x)_\# \boldsymbol{\lambda}_{g^i_x}.$$

Hence

$$J_{g^0_x} \cdot (L_x)_\# \boldsymbol{\lambda}^0_x = J^1_x \cdot (L_x)_\# \boldsymbol{\lambda}_{g^1_x} = \big(\, J^1_x \sqrt{\det S_x} \,\big) \cdot (L_x)_\# \boldsymbol{\lambda}_{g^0_x}$$

so that

$$J^1_x = \frac{1}{\sqrt{\det S_x}} J^0_x.$$

Hence

$$\mathbb{E}\big[\, J_x^1 \big|\, \Psi(x) = 0\,\big]\, \mathrm{vol}_{g^1}\big[\, dx\,\big] = \frac{1}{\sqrt{\det S_x}} \mathbb{E}\big[\, J_x^0 \big|\, \Psi(x) = 0\,\big] \sqrt{\det S_x}\, \mathrm{vol}_{g_0}\big[\, dx\,\big]$$

proving that the right-hand side of (2.2.16) is independent of the metric g.

3. Independence of the metric h. Let h^0, h^1 be two metrics on E. We denote by $\mathrm{Var}^i(x)$ the variance operator of $\Psi(x)$ determined by the inner product $h^i(x)$ on E_x. The probability *density* of $\Psi(x)$ depends on the inner product on E_x. We denote by $p^i_{\Psi(x)}$ the probability density of $\Psi(x)$ determined by the inner product $h^i(x)$. We have

$$p^i_{\Psi(x)}(0) = \frac{1}{\sqrt{\det\big(\, 2\pi\, \mathrm{Var}^i(x)\,\big)}}.$$

There exists a smooth endomorphisms G of E which is symmetric and positive definite with respect to h^0 and such that

$$h^1(u,v) = h^0\big(\, Gu, v\,\big)$$

for any smooth sections u, v of E. As explained in Remark 1.1.17 we have

$$\mathrm{Var}^1(x) = \mathrm{Var}^0(x) G(x)$$

so that

$$p^1_{\Psi(x)}(0) = \frac{1}{\sqrt{\det G(x)}} p^0_{\Psi(x)}.$$

Using the same notations as in the proof of the independence of g we have

$$\boldsymbol{\lambda}_{h_x^i} = J_x^i \cdot (L_x)_\# \boldsymbol{\lambda}_{g_x},\ \ i = 0, 1$$

and

$$\boldsymbol{\lambda}_{h_x^1} = \sqrt{\det G(x)}\boldsymbol{\lambda}_{h_x^0}$$

from which we deduce that $J_x^1 = \sqrt{\det G(x)} \cdot J_x^0$. Hence

$$\mathbb{E}\big[\, J_x^1 \big|\, \Psi(x) = 0\,\big] p^1_{\Psi(x)}(0) = \sqrt{\det G(x)} \mathbb{E}\big[\, J_x^0 \big|\, \Psi(x) = 0\,\big] \frac{1}{\sqrt{G(x)}} p^0_{\Psi(x)}(0).$$

The point of the above exercise is clear: if we could prove (2.2.16) for *some* convenient choices of g, h, ∇^E, then we would have a proof for *any* choices of g, h, ∇^E.

Using partitions of unity we can reduce (2.2.16) to the case when φ is supported on a open subset $\mathcal{O} \subset M$ with the following properties.

- The open subset $\mathcal{O}$ is diffeomorphic to an open subset $\mathscr{V} \subset \mathbb{R}^m$.
- The restriction of E to $\mathcal{O}$ is trivializable. Fix one such trivialization, $E|_{\mathcal{O}} \cong \underline{\mathbb{R}}^m_{\mathcal{O}}$. Here $\underline{\mathbb{R}}^m_{\mathcal{O}}$ denote the trivial vector bundle over $\mathcal{O}$ with fiber $\mathbb{R}^m$.

Then we can identify the restriction to $\mathcal{O}$ of the section Ψ with a random Gaussian map $\Psi : \mathscr{V} \to \mathbb{R}^m$. Suppose now that the restriction of g to $\mathcal{O}$ corresponds to the Euclidean metric on $\mathscr{V}$, the restriction of h to $E|_{\mathcal{O}}$ corresponds to the trivial metric on the trivial bundle $\underline{\mathbb{R}}^m_{\mathscr{V}}$, and the restriction of ∇^E to $E|_{\mathcal{O}}$ corresponds to the trivial connection on $\underline{\mathbb{R}}^m_{\mathscr{V}}$.

In this case (2.2.16) reduces to the local Kac-Rice formula (2.2.9).

□

Remark 2.2.17. The equality (2.2.16) displays a remarkable phenomenon: the quantity

$$\mathbb{E}\big[\, J_{\nabla^E \Psi(x)} \big|\, \Psi(x) = 0\,\big] p_{\Psi(x)(0)} \varphi(x) \operatorname{vol}_g \big[\, dx\,\big]$$

is independent of the various geometric choices: the metrics on M and E and the connection on E. It depends only on the ample Gaussian C^1 section Ψ. Since this quantity is something one can integrate over a manifold it is called a 1-*density* in geometric parlance; see [**117**, Sec.3.4]. Thus the global Kac-Rice formula describes explicitly a canonical 1-density on M whose integral over a Borel subset gives the expected number of zeros of ψ in that Borel subset. We will refer to it as the *Kac-Rice* 1-*density*.

☞ *A word of warning!* The concept of 1-density is not to be confused with the concept of density used in analysis and in physics. For example, a 1-density on a Riemann manifold (M, g) is essentially a measure ν on M that is absolutely continuous with respect to the volume measure vol_g determined by the metric g.

The analysts' and physicists' density is the Radon-Nicodym derivative $\frac{d\nu}{d\operatorname{vol}_g}$, also called the density of ν relative to vol_g. When working on $\mathbb{R}^n$ with the usual Euclidean metric, then $d\operatorname{vol}_g$ is the Lebesgue measure the concepts of 1-density and density tend to be confused. In geometry this confusion could lead to erroneous conclusions.

We could have formulated the Kac-Rice formula in the language of 1-densities from the start. We chose not to do so since the concept of 1-density is not widely known and can obscure the simple nature of this result in concrete situations. □

Let us mention a few immediate consequences of the global Kac-Rice formula.

Corollary 2.2.18. *Suppose that Ψ is an ample Gaussian C^1-section of the smooth vector bundle $E \to M$ of rank $m = \dim M$. Then, for any compact set $K \subset M$, the expected number of zeros of Ψ inside K is finite.* □

Corollary 2.2.19. *Suppose that (M, g) is a compact Riemannian manifold and $F : M \to \mathbb{R}$ is a Gaussian C^2 function such that the random section $dF : M \to T^*M$ is ample. Denote by ∇^g the Levi-Civita connection on T^*M. Then the following hold.*

(i) *The function F is almost surely a Morse function, i.e., all its critical points are nondegenerate.*

(ii) *For any continuous function $\varphi : M \to \mathbb{R}$ we have*

$$\mathbb{E}\big[\, Z(\varphi, dF)\,\big] = \int_M \varphi(x) \mathbb{E}\big[\, \operatorname{Hess}_F(x) \big|\, df(x) = 0\,\big] p_{df(x)}(0) \operatorname{vol}_g \big[\, dx\,\big],$$

where the $\operatorname{Hess}_F(x)$ is Hessian of f at x, $\operatorname{Hess}_F(x) = \nabla^g dF(x)$.

□

Remark 2.2.20. In the above corollary we need not have fixed a metric. As explained in Subsection 1.2.7, the Gaussian function F defines both a metric on T^*M (hence a metric g^{corr} on M) and a connection on T^*M compatible with it. This connection is the Levi-Civita connection of the metric g^{corr}. We could have described the Kac-Rice 1-density entirely in terms of g^{corr}. □

Similarly, Corollary 2.2.15 has a global counterpart.

Corollary 2.2.21. *Suppose that* (M, g) *is a compact Riemannian manifold and* $F : M \to \mathbb{R}$ *is a* J_1*-ample* C^2 *Gaussian function, i.e.,*

$$\text{for any } \boldsymbol{p} \in M \text{ the Gaussian vector } F(\boldsymbol{p}) \oplus dF(\boldsymbol{p}) \in \mathbb{R} \oplus T^*_{\boldsymbol{p}}M \text{ is nondegenerate.}$$

Denote by $\boldsymbol{D}_F$ *the discriminant measure of* F

$$\boldsymbol{D}_F = \sum_{dF(\boldsymbol{p})=0} \delta_{F(\boldsymbol{p})} = \sum_{t \in \mathbb{R}} \#\big\{\, \boldsymbol{p}; F(\boldsymbol{p}) = t,\;\; dF(\boldsymbol{p}) = 0 \,\big\} \delta_t.$$

Then F *is* a.s. *Morse and, for any* $\varphi \in C^0_{\mathrm{cpt}}(\mathbb{R})$ *we have*

$$\begin{gathered}\mathbb{E}\big[\, \boldsymbol{D}_F[\varphi] \,\big] \\ = \int_M \left(\int_{\mathbb{R}} \mathbb{E}\big[\, |\det \mathrm{Hess}_F(x)| \,\big|\, dF(x) = 0, F(x) = t \,\big] \varphi(t) \mathbb{P}_{F(x)}\big[\, dt \,\big] \right) p_{dF(x)}(0) \operatorname{vol}_g\big[\, dx \,\big].\end{gathered} \tag{2.2.17}$$

□

Note that both sides of the above equality are Daniell integrals. The Daniel-Stone theorem [**49**, Thm. 4.5.2] imples that (2.2.17) holds with φ replaced by an arbitrary bounded measurable function.

2.3. Applications

In the immortal words of Yogi Berra "*in theory there is no difference between theory and practice. In practice there is.*" The applications of the Kac-Rice formula are good illustrations of the above principle. It this section we will show how the Kac-Rice formula works in some concrete situations. We start with the 1-dimensional situation. Even in this simplest of the situations we will reach beautiful geometric conclusions.

2.3.1. Some one-dimensional applications. Suppose that $I \subset \mathbb{R}$ is an open interval of the real axis and $F : I \to \mathbb{R}$ is a centered C^1 Gaussian function such that for any $t \in I$ the Gaussian random variable $F(t)$ is nondegenerate. Let $K : I \times I \to \mathbb{R}$ be the covariance kernel of $F(t)$, i.e.,

$$K(t, s) = \mathbb{E}\big[\, F(t)F(s) \,\big], \;\; \forall s, t \in I.$$

Since $F(t)$ is nondegenerate we have $K(t, t) > 0$, $\forall t$. We set $Z[F] := Z[I, F]$, i.e., $Z[F]$ is the number of zeros of F in I.

The local Kac-Rice formula implies that

$$\mathbb{E}\big[\, Z[F] \,\big] = \int_I \underbrace{\mathbb{E}\big[\, |F'(t)| \,\big|\, F(t) = 0 \,\big] p_{F(t)}(0)}_{=:\rho_t} \, dt. \tag{2.3.1}$$

We need to clarify the nature of the integrand ρ_t in the above equality. Observe first that

$$p_{F(t)}(0) = \frac{1}{\sqrt{2\pi K(t, t)}}.$$

Observe next that $F'(t)$ is a continuous Gaussian function with covariance kernel

$$\mathbb{E}\big[\, F'(t)F'(s) \,\big] = \partial^2_{ts} K(t, s).$$

Note also that

$$\mathbb{E}\big[\,F'(t)F(t)\,\big] = \partial_t K(t,s)_{s=t}.$$

The Gaussian regression formula (1.1.20) shows that

$$\mathbb{E}\big[\,|F'(t)|\,\big|F(t)=0\,\big] = \mathbb{E}\big[\,\big|\,X\,\big|\,\big]$$

where X is a centered Gaussian random variable with variance,

$$v_t = \partial^2_{ts}K(t,s)_{s=t} - \frac{\partial_t K(t,s)^2_{s=t}}{K(t,t)}.$$

Hence

$$\mathbb{E}\big[\,|F'(t)|\,\big|F(t)=0\,\big] \stackrel{(1.1.8)}{=} (2v_t/\pi)^{1/2},$$

and

$$\begin{gathered}\rho_t = \left(\frac{\partial^2_{st}K(t,s)K(t,s) - \partial_t K(t,s)^2}{K(t,s)^2}\right)^{1/2}_{s=t} = \sqrt{\partial^2_{s,t}\log K(t,s)_{t=s}},\\ \mathbb{E}\big[\,Z[F]\,\big] = \frac{1}{\pi}\int_I \rho_t dt.\end{gathered} \tag{2.3.2}$$

Example 2.3.1 (Kac polynomials)**.** Suppose that $F(t)$ is a random polynomial of the form

$$F(t) = F_n(t) = \sum_{k=0}^n A_k t^k,$$

where the coefficients are independent standard normal random variables. In this case the covariance kernel is

$$K(s,t) = \frac{1-(st)^{n+1}}{1-st}.$$

Denote by Z_n the numbers of real roots on F_n.

Such random polynomials are referred to as *Kac polynomials*, the Kac in Kac-Rice. They were first considered by M. Kac [**77**] in 1943 when he proved the first version of (2.3.2). More precisely he showed that

$$\mathbb{E}\big[\,Z_n\,\big] = \frac{1}{\pi}\int_{\mathbb{R}} \sqrt{f_n(t)}\,dt,\;\; f_n(t) := \frac{1}{(t^2-1)^2} - \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}. \tag{2.3.3}$$

For example,

$$f_2(t) = \frac{1}{(t^2-1)^2} - \frac{9t^4}{(t^6-1)^2} = \frac{t^4+t^2+1}{(t^4+t^2+1)^2},$$

and

$$\mathbb{E}\big[\,Z_n\,\big] \approx 0.5055.$$

In particular, we deduce that

$$\mathbb{P}\big[\,Z_2>0\,\big] = \frac{1}{2}\mathbb{E}\big[\,Z_2\,\big] \approx 0.25.$$

In [**77**] M. Kac proved the rather surprising result

$$\mathbb{E}\big[\,Z_n\,\big] = \frac{2}{\pi}\log N + O(1), \;\; \text{as } n\to\infty.$$

This can be a bit refined; see [51, Sec.2.5]. More precisely, there exists a universal constant $C > 0$ ($C \approx 0.6257...$) such that as $n \to \infty$ we have

$$Z_n = \frac{2}{\pi}\left(\log n + C + \frac{2}{n\pi} \right) + O\big(1/n^2\big). \tag{2.3.4}$$

The results in [44] imply that the expected number of critical points of $F_n(t)$ is also of the order $\log n$ as $n \to \infty$. □

Example 2.3.2 (The Kostlan statistics). Consider the Kostlan random polynomials

$$F_n(t) = \sum_{k=0}^{n} X_k t^k,$$

where the coefficients X_k are independent normal random variables with mean zero and variances

$$\operatorname{Var}\big[\, X_k \,\big] = \binom{n}{k}.$$

Denote by Z_n the number of real zeros of F_n. In this case the covariance kernel is

$$K(s,t) = \sum_{k=0}^{n} \binom{n}{k}(st) = (1+st)^N$$

and we have

$$\log K(s,t) = N\log(1+st),\ \ \partial_t \log K(s,t) = \frac{Ns}{1+st},$$

$$\partial^2_{st} \log K(s,t) = \frac{N}{(1+st)^2},\ \ \rho_t = \sqrt{\partial^2_{st}\log K(s,t)|_{s=t}} = \frac{\sqrt{N}}{1+t^2}.$$

The Kac-Rice formula implies that the expected number of zeros is

$$\mathbb{E}\big[\, Z_n \,\big] = \frac{2\sqrt{N}}{\pi}\int_0^\infty \frac{1}{1+t^2} = \sqrt{n}.$$

We see that the Kostlan random polynomials have, on average, more real zeros than the Kac random polynomials. □

Example 2.3.3 (The Legendre statistics). Recall that the Legendre polynomials are obtained from the sequence of monomials $(t^k)_{k\geq 0}$ by applying the Gramm-Schmidt procedure with respect to the inner product in $L^2([-1,1],dt)$.

Concretely, the degree n Legendre polynomial is

$$p_n(t) := \sqrt{\frac{2n+1}{2}}\,\ell_n(t),\ \ \ell_n(t) := \frac{1}{2^n n!}\frac{d^n}{dt^n}\big(t^2-1\big)^n. \tag{2.3.5}$$

We can construct a random polynomial

$$F_N(t) = \sum_{k=0}^{N} X_k p_k(t),$$

where X_k are independent standard normal random variables, $\forall k$. Using the Christoffel-Darboux theorem [149] we deduce that its covariance kernel is given by

$$K_N(s,t) = \sum_{k=0}^{N} p_k(s)p_k(t) = \frac{N+1}{2} \cdot \frac{\ell_{N+1}(t)\ell_N(s) - \ell_{N+1}(s)\ell_N(t)}{t-s}. \qquad \square$$

M. Das [42] has shown that the expected number of zeros of $F_N(t)$ in $[-1,1]$ is asymptotic to $\frac{1}{\sqrt{3}}N$ for large N. The Legendre ensemble displays an even stronger bias towards a relatively large number of real roots. The reason is that the number of zeros of the Legendre polynomial ℓ_n goes to ∞ as $n \to \infty$. $\square$

I want to describe two nice geometric applications of the 1-dimensional Kac-Rice formula. My presentation follows [119].

Example 2.3.4. Suppose that C is a smooth closed curve on the unit n-dimensional sphere

$$S^n = \left\{ (x_0, x_1, \dots, x_n) \in \mathbb{R}^{n+1};\;\; x_0^2 + x_1^2 + \cdots + x_n^2 = 1 \right\}.$$

Denote by L its length. By Equator in S^n we mean an $(n-1)$-dimensional sphere obtained by intersecting S^n with a hyperplane through the origin. An Equator divides the sphere into two parts called hemispheres. Using the Kac-Rice formula we will prove, in one stroke, two related facts.

(i) If $L < 2\pi$, then this curve is entirely contain in some hemisphere.

(ii) If $L > 2\pi$, then there exists an Equator of the sphere that intersects the curve in at least four points.

The case $n=2$ of (i) seems to be part of the folklore of mathematics; see e.g. [153, Problem 1.10.4]. The case $n=2$ of (ii) was proved more recently, in a 2008 American Mathematical Monthly paper, [72]. The authors refer to it as a 1969 conjecture of Hugo Steinhaus. Here is a probabilistic proof of these facts.

Parametrize C by arclength, $[0,L] \ni s \mapsto \boldsymbol{x}(s) := \big(x_0(s), \dots, x_n(s)\big) \in \mathbb{R}^{n+1}$. Since $C \subset S^n$ we have $|\boldsymbol{x}(s)| = 1$, $\forall s$, where $|-|$ denotes the natural Euclidean norm. Moreover, since this is arclength parametrization we have $|\boldsymbol{x}'(s)| = 1$, $\forall s$.

Any vector $\boldsymbol{u} \in \mathbb{R}^{n+1}$ determines a linear functional $\ell_{\boldsymbol{u}} : \mathbb{R}^{n+1} \to \mathbb{R}$, $\ell_{\boldsymbol{u}}(\boldsymbol{x}) = \langle \boldsymbol{u}, \boldsymbol{x} \rangle$, where $\langle -,- \rangle$ is the canonical inner product in $\mathbb{R}^{n+1}$. To prove (i), we have to show that there exists $\boldsymbol{u} \neq 0$ such that the restriction of $\ell_{\boldsymbol{u}}$ to C has no zeros. To prove (ii), we have to show that there exists $\boldsymbol{u} \neq 0$ such that the restriction of $\ell_{\boldsymbol{u}}$ to C has at least four zeros.

The restriction of $\ell_{\boldsymbol{u}}$ to C can be identified with the function $f_{\boldsymbol{u}} : [0,L] \to \mathbb{R}$, $f_{\boldsymbol{u}}(s) = \langle \boldsymbol{u}, \boldsymbol{x}(s) \rangle$. Choose independent standard random variables $(U_k)_{0 \le k \le n}$ and form the random Gaussian function

$$F_U : [0,L] \to \mathbb{R}, \;\; F_U(s) = \sum_{i=0}^{n} U_i x_i(s).$$

Its covariance kernel is $K(s,t) = \langle \boldsymbol{x}(s), \boldsymbol{x}(t) \rangle$. We deduce

$$a_t = \langle \boldsymbol{x}(t), \boldsymbol{x}(t) \rangle = |\boldsymbol{x}(t)|^2 = 1, \;\; b_t = K'_t(s,t)\big|_{s=t} = \langle \boldsymbol{x}(t), \boldsymbol{x}'(t) \rangle = \frac{1}{2}\frac{d}{dt}|\boldsymbol{x}(t)|^2 = 1,$$

$$c_t = K''_{st}(s,t)\big|_{s=t} = \big|\, \boldsymbol{x}'(t)\, \big|^2 = 1.$$

The ampleness of F_U follows from the equality $a_t = 1$. We can apply the Kac-Rice formula to deduce that the expected number of zeros of F_U is $Z_C = \frac{L}{\pi}$.

To reach the conclusions (i) and (ii) we need an additional input, topological in nature. Observe that if $f_{\boldsymbol{u}}$ has only nondegenerate zeros, then it has an even number of them. Indeed, a nondegenerate zero of $f_{\boldsymbol{u}}$ corresponds to a point where the curve C crosses the hyperplane $\{\ell_{\boldsymbol{u}} = 0\}$ transversally from one side to the other. Since the curve is closed, it must cross this hyperplane an even number number of times.

The ampleness condition $a_t > 0$ implies that the zeros of F_U are almost surely nondegenerate. Thus, almost surely, the function F_U has an even number of zeros.

If $L < 2\pi$, then $Z_C < 2$, and the probability that the number of zeros of F_U is < 2 is positive. Since F_U has an even number of zeros we deduce that the probability that F_U has no zeros is positive. This proves (i).

If $L > 2\pi$, then $Z_C > 2$. Hence, the probability that F_U has more than two zeros is positive and we deduce that the probability that F_U has at least four zeros is positive. This proves (ii). □

Example 2.3.5 (Fáry-Milnor)**.** Suppose that

$$[0, L] \ni s \mapsto \boldsymbol{r}(s) = (x(s), y(s), z(s)) \in \mathbb{R}^3$$

is the arclength parametrization of a smooth knot K in $\mathbb{R}^3$. Here, by knot I understand a smoothly embedded S^1 in $\mathbb{R}^3$.

Consider a random linear function

$$H : \mathbb{R}^3 \to \mathbb{R}, \;\; H(x, y, z) = Ax + By + Cz,$$

where A, B, C are independent standard normal random variables with mean zero and variance 1. Denote by $\mu(H) = \mu(A, B, C)$ the number of critical points of the restriction of H to the knot. These are the points on the knot where the vector (A, B, C) is perpendicular to the tangent vector to the curve at that point.

The restriction of H to the K is described the Gaussian random function

$$F(s) = Ax(s) + By(s) + Cz(s).$$

Note that the critical points of $H|_K$ correspond to the zeros of the derivative.

$$F'(s) = Ax'(s) + By'(s) + Cz'(s)$$

The derivative $F'(s)$ is a Gaussian random function with covariance kernel

$$K(s_1, s_2) = x'(s_1)x'(s_2) + y'(s_1)y'(s_2) + z'(s_1)z'(s_2) = \boldsymbol{T}(s_1) \bullet \boldsymbol{T}(s_2),$$

where $(\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B})$ is the Frénet frame along the curve and $\bullet$ denotes the standard inner/dot product in $\mathbb{R}^3$. We have

$$\partial_{s_2} K(s_1, s_2) = \boldsymbol{T}(s_1) \bullet \boldsymbol{T}'(s_2) = \kappa(s_2)\boldsymbol{T}(s_1) \bullet \boldsymbol{N}(s_2),$$

where κ denotes the curvature of the curve. Similarly

$$\partial^2_{s_1 s_2} K(s_1 s_2) = \kappa(s_1)\kappa(s_2)\boldsymbol{N}(s_1) \bullet \boldsymbol{N}(s_2).$$

We deduce

$$K(s, s) = 1, \;\; \partial_{s_2} K(s, s) = 0, \;\; \partial^2_{s_1 s_2} K(s, s) = \kappa(s)^2, \;\; \rho_s = |\kappa(s)|.$$

Hence the ampleness assumption is satisfied. The Kac-Rice formula implies that the critical points of $H|_K$ are almost surely nondegenerate. They come in two types: local minima and local maxima. We denote by $m_\pm(H)$ the number of local minima/maxima of $H|_K$. Then, almost surely, $m_-(H) = m_+(H)$ and $\mu(H) = m_-(H) + m_+(H)$. The Kac-Rice formula (2.3.2) implies that

$$\mathbb{E}\big[\,\mu(H)\,\big] = 2\mathbb{E}\big[\,m_+(H)\,\big] = \frac{1}{\pi}\int_0^L |\kappa(s)|ds = \frac{1}{\pi}\times \text{the total curvature of the curve.}$$

This result was first proved independently by I. Fáry [**56**] and J. Milnor [**103**]. In particular, Milnor, who was an undergraduate at the time, used this to prove a conjecture of K. Borsuk roughly stating that, to knot a curve, you need to bend it quite a bit. More precisely, if the total curvature is $\leq 4\pi$, then the knot K is the unknot. □

2.3.2. The distribution of critical points. Things get substantially more involved for Gaussian functions of several variables, but in certain cases we can still say something that is meaningful.

Example 2.3.6 (Random trigonometric polynomials with given Netwon polyhedron)**.** Consider the random Gaussian trigonometric polynomial

$$X_N(\vec{\theta}) = X_N(\theta_1,\dots,\theta_m)$$

defined in Example 1.2.17. Its covariance kernel is

$$\mathcal{K}(\vec{\theta},\vec{\phi}) = S_N\big(\,\vec{\theta}-\vec{\phi}\,\big) = \sum_{\vec{\ell}\in P_N} e^{\boldsymbol{i}\langle\vec{\ell},\vec{\theta}-\vec{\phi}\rangle}.$$

Above $P_N = N\cdot P$, where P is a fixed Newton polyhedron, i.e., a convex polyhedron in $\mathbb{R}^m$, with vertices lattice points, containing the origin in the interior and symmetric about the origin. For $m=1$ and $P=[-1,1]$ the number of zeros of X_N were first investigated by S. O. Rice [**135**], the Rice in Kac-Rice.

Denote by Z_N or $Z(P_N)$ the number of critical points of X_N on the torus $\mathbb{T}^m = \big(\mathbb{R}/2\pi\mathbb{Z}\big)^m$, or equivalently, the number of zeros of its gradient $G_N = \nabla X_N$ on the box $B_m := [0,2\pi]^m$. We have seen in Example 1.2.17 that G_N is ample for all N sufficiently large. We deduce from the Kac-Rice formula (**KR**) that

$$\mathbb{E}\big[\,Z_N\,\big] = \int_{B_m}\mathbb{E}\big[\,|\det\operatorname{Hess}_{X_N}(\vec{\theta})|\;\big|\;\nabla X_N(\vec{\theta})=0\,\big]p_{\nabla X_N(\theta)}(0)d\vec{\theta}$$

The computations in Example 1.2.17 show that the above integrand is independent of $\vec{\theta}$. We have

$$\begin{aligned} p_{\nabla X_N(\theta)}(0) = \frac{1}{\sqrt{\det(2\pi)\operatorname{Var}\big[\,\nabla X_N(0)\,\big]}} &\overset{(1.2.12)}{\sim} \frac{1}{\sqrt{\det(2\pi)N^{m+2}M(P)}} \\ &\sim (2\pi)^{-m/2}\big(\,\det M(P)\,\big)^{-1/2}N^{-m(m+2)/2}\quad N\to\infty. \end{aligned} \tag{2.3.6}$$

The Hessian $H_N(\vec{\theta}) = \operatorname{Hess}_{X_N}(\vec{\theta})$ is a Gaussian vector valued in the space $\mathbf{Sym}(\mathbb{R}^m)$ of symmetric $m\times m$-matrices. This is a Euclidean space of dimension $m(m+1)/2$ with inner product

$$\langle A,B\rangle = \operatorname{tr}(AB).$$

On $\mathbf{Sym}(\mathbb{R}^n)$ we can use the *orthonormal* coordinates $\big(\omega_{ij}\big)_{1\leq i\leq j\leq m}$, where

$$\omega_{ij}(A)=\begin{cases}A_{ii}, & i=j,\\ \sqrt{2}A_{ij}, & i\neq j.\end{cases}$$

The variance operator $\operatorname{Var}\big[\,H_N(\vec{\theta})\,\big]$ is given by the matrix

$$Q(N)=\big(\,Q(N)_{ij|k\ell}\,\big)_{\substack{i\leq i\leq j\leq m,\\ 1\leq k\leq \ell\leq m}}$$

where

$$Q(N)_{ij|k\ell}=\mathbb{E}\big[\,\omega_{ij}(H_N)\omega_{k\ell}(H_N)\,\big]$$

Consider the quadratic functions $q_{ij}:P\to\mathbb{R}$

$$q_{ij}(x_1,\dots,x_m)=\begin{cases}x_i^2, & i=j\\ \sqrt{2}x_ix_j, & i\neq j.\end{cases}$$

We denote by $Q(P)$ the Gramian matrix of the collection $\big(\,q_{ij}\,\big)_{1\leq i\leq j\leq m}$ with respect to the inner product on $L^2(P,\boldsymbol{\lambda})$. This is a nondegenerate symmetric matrix of size $\frac{m(m+1)}{2}\times\frac{m(m+1)}{2}$. From the equalities

$$\mathbb{E}\big[\,H_N(\vec{\theta})_{ij}H_N(\vec{\theta})_{k\ell}\,\big]=\partial^2_{\theta_i\theta_j}\partial^2_{\phi_k\phi_\ell}S_N(\vec{\theta}-\vec{\phi}\,)_{\vec{\theta}=\vec{\phi}}=\partial^4_{\tau_i\tau_j\tau_k\tau_\ell}S_N(0)$$

and (1.2.7) we deduce that

$$Q(N)\sim N^{m+4}Q(P)\big(\,1+O(1/N)\,\big),$$

To compute the conditional expectation $\mathbb{E}\big[\,|\det H_N(\vec{\theta})|\big|\,\nabla X_N(\theta)=0\,\big]$ I plan to use the Gaussian regression formula so we need to find the correlation operator $C_{H_N(\vec{\theta}),\nabla X_N(\vec{\theta})}$. Note that

$$\mathbb{E}\big[\,H_N(\vec{\theta})_{ij}\partial_{\theta_k}X_N(\vec{\theta})\,\big]=\partial^2_{\theta_i\theta_j}\partial_{\phi_k}S_N(\vec{\theta}-\vec{\phi}\,)=\boldsymbol{i}\partial^3_{\tau_i\tau_j\tau_k}S_N(0)=0.$$

Hence the Gaussian vectors $H_B(\vec{\theta})$ and $\nabla X_N(\vec{\theta})$ are independent. Denote by $\Gamma_{Q(N)}$ the centered Gaussian measure on $\mathbf{Sym}(\mathbb{R}^n)$ with variance $Q(N)$. We deduce

$$\mathbb{E}\big[\,|\det H_N(\vec{\theta})|\,\big|\,\nabla X_N(\theta)=0\,\big]=\int_{\mathbf{Sym}(\mathbb{R}^m)}\big|\,\det A\,\big|\Gamma_{Q(N)}\big[\,dA\,\big]$$

$(A=N^{(m+4)/2}B)$

$$=N^{m(m+4)/2}\int_{\mathbf{Sym}(\mathbb{R}^m)}\big|\,\det B\,\big|\Gamma_{N^{-(m+4)}Q(N)}\big[\,dB\,\big].$$

Observing that $N^{-(m+4)}Q(N)\sim Q(P)$ as $N\to\infty$ we deduce

$$\mathbb{E}\big[\,|\det H_N(\vec{\theta})|\,\|\,\nabla X_N(\theta)=0\,\big]\sim N^{m(m+4)/2}\int_{\mathbf{Sym}(\mathbb{R}^m)}\big|\,\det B\,\big|\Gamma_{Q(P)}\big[\,dB\,\big].$$

The last integral is strictly positive since the Gaussian random matrix $Q(P)$ is nondegenerate.

Putting together all of the above we deduce that as $N\to\infty$ we have

$$\mathbb{E}\big[\,Z_N\,\big]\sim\operatorname{vol}(B_m)\times(2\pi)^{-m/2}\big(\,\det M(P)\,\big)^{-1/2}N^{-m(m+2)/2}$$
$$\times N^{m(m+4)/2}\int_{\mathbf{Sym}(\mathbb{R}^m)}\big|\,\det B\,\big|\,\Gamma_{Q(P)}\big[\,dB\,\big]$$

$$\sim (2\pi)^{m/2}\big(\det M(P)\big)^{-1/2}\left(\int_{\mathbf{Sym}(\mathbb{R}^m)} |\det B|\,\Gamma_{Q(P)}\big[\,dB\,\big]\right) N^m$$

$$\sim (2\pi)^{m/2}\big(\det M(P)\big)^{-1/2}\left(\int_{\mathbf{Sym}(\mathbb{R}^m)} |\det B|\,\Gamma_{Q(P)}\big[\,dB\,\big]\right)\frac{\mathrm{vol}(P_N)}{\mathrm{vol}(P)}$$

$$= K(P)\,\mathrm{vol}(P_N),$$

where

$$K(P) = \frac{(2\pi)^{m/2}}{\big(\det M(P)\big)^{1/2}\mathrm{vol}(P)}\left(\int_{\mathbf{Sym}(\mathbb{R}^m)} |\det B|\,\Gamma_{Q(P)}\big[\,dB\,\big]\right)$$

The results of Bernshtein [**18**] and Kouchnirenko [**83**] imply the *very rough* upper bound

$$Z\big(P_N\big) \leq m!\,\mathrm{vol}\,\big(P_N\big)\ \text{ a.s.} \tag{2.3.7}$$

The result we have just proved shows that as $N\to\infty$, $\mathbb{E}\big[\,Z_N\,\big]$ has the same order of growth as $\mathrm{vol}\,\big(P_N\big)$ indicating that the mean of Z_N is close its theoretical max.

In [**110**] I computed the constant $K(P)$ for various polyhedra P. As a special case let me mention the polygon in $\mathbb{R}^2$ with six vertices

$$(1,1),(-1,-1),(1,0),(-1,0),(0,1),(0,-1).$$

Its area is 3, so the inequality (2.3.7) predicts that $Z(P)\leq 6$ a.s..

V.I. Arnold has proved in [**8**] that this inequality is sharp, i.e., $\sup Z(P) = 6$ a.s.. He achieved this using topological techniques that allowed him to conclude that there exists a trigonometric polynomial with Newton polygon P which is Morse and has exactly 6 critical points.

In [**110**] I proved that

$$\mathbb{E}\big[\,Z(P)\,\big] = \frac{4\pi}{3} \approx 4.188$$

Hence, with positive probability, there must exist trigonometric polynomials with Newton polygon P and at least 5 critical points. The random trigonometric polynomial $X_1(P)$ is a.s. Morse so it has an even number of critical points almost surely. Hence, the probability that there exists a Morse polynomial $X_1(P)$ with exactly 6 critical points is positive. □

Example 2.3.7 (Isotropic Gaussian functions)**.** This example might seem rather special, but as we will see later on, it is universal.

Consider the smooth isotropic Gaussian function $\Phi=\Phi_{\mathfrak{a}}$ on $\mathbb{R}^m$ constructed in Example 1.2.35. We recall its construction.

We fix an amplitude $\mathfrak{a}$, i.e., an even Schwartz function $\mathfrak{a}\in\mathcal{S}\big(\mathbb{R}\big)$ such that $\mathfrak{a}(0)=1$. Consider the finite measure $\mu_{\mathfrak{a}}\in\mathrm{Meas}\,\big(\mathbb{R}^m\big)$

$$\mu_{\mathfrak{a}}\big[\,d\xi\,\big] = \frac{1}{(2\pi)^m}w_{\mathfrak{a}}(\xi)\boldsymbol{\lambda}\big[\,d\xi\,\big],\ \ w_{\mathfrak{a}}(\xi) = w_{\mathfrak{a},m}(\xi) = \mathfrak{a}\big(\,|\xi|\,\big)^2.$$

The statistic of the random function $\Phi_{\mathfrak{a}}$ is determined by its covariance kernel

$$\mathcal{K}_{\mathfrak{a}}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{x}-\boldsymbol{y}\,\big),\ \ \boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{x}\,\big) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}w_{\mathfrak{a}}(\xi)d\xi \tag{2.3.8}$$

The Gaussian function $\Phi_{\mathfrak{a}}$ is a.s. smooth, isotropic and $\mu_{\mathfrak{a}}$ is its spectral measure. We set

$$\mathfrak{C}_{\mathfrak{a}} := \mathfrak{C}[-, \Phi_{\mathfrak{a}}] = \sum_{\nabla \Phi_{\mathfrak{a}}(\boldsymbol{x})=0} \delta_{\boldsymbol{x}}. \tag{2.3.9}$$

Thus, $\mathfrak{C}_{\mathfrak{a}}[B]$ is the number of critical points of $\Phi_{\mathfrak{a}}$ in B. I want to compute $\mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[B]\,\big]$ for a box $B \subset \mathbb{R}^m$.

For any multi-indices $\alpha, \beta \in \big(\mathbb{Z}_{\geq 0}\big)^m$ we have

$$\begin{aligned}\mathbb{E}\big[\,\partial^\alpha \Phi_{\mathfrak{a}}(\boldsymbol{x}) \partial^\beta \Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]_{\boldsymbol{x}=\boldsymbol{y}} &= \partial^\alpha_{\boldsymbol{x}} \partial^\beta_{\boldsymbol{y}} \mathcal{K}_{\mathfrak{a}}(\boldsymbol{x}, \boldsymbol{y})\big|_{\,\boldsymbol{x}=\boldsymbol{y}} \\ &= \frac{(-1)^{|\beta|} \boldsymbol{i}^{|\alpha|+|\beta|}}{(2\pi)^m} \int_{\mathbb{R}^m} \xi^{\alpha+\beta} \mathfrak{a}(|\xi|)^2 d\xi.\end{aligned} \tag{2.3.10}$$

For any multi-index $\alpha \in \big(\mathbb{Z}_{\geq 0}\big)^m$ we set

$$M^\alpha_{\mathfrak{a}} := \int_{\mathbb{R}^m} \xi^\alpha \mu_{\mathfrak{a}}\big[\,d\xi\,\big] = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} \xi^\alpha \mathfrak{a}(|\xi|)^2 d\xi.$$

We say that the multi-index $\alpha = (\alpha_1, \dots, \alpha_m)$ is *even* if α_j is even for any $j = 1, \dots, m$. The multi-index α is called *odd* if it is not even. The radial symmetry of $\mathfrak{a}\big(\,|\xi|\,\big)$ implies that

$$M^\alpha_{\mathfrak{a}} = 0 \;\text{ if } \alpha \text{ is odd.} \tag{2.3.11}$$

Using spherical coordinates on $\mathbb{R}^m$ we deduce that for any α we have

$$M^\alpha_{\mathfrak{a}} = \frac{1}{(2\pi)^m} \left(\int_0^\infty r^{m-1+|\alpha|} \mathfrak{a}(r)^2 dr \right) \times \underbrace{\int_{S^{m-1}} \xi^\alpha \,\mathrm{vol}_{S^{m-1}}\big[\,d\xi\,\big]}_{=:\boldsymbol{m}_\alpha}. \tag{2.3.12}$$

Note that $\boldsymbol{m}_\alpha$ is independent of $\mathfrak{a}$. In particular, $\boldsymbol{m}_0$ is the "area" off the unit sphere S^{m-1}.

If we let $\mathfrak{a}_0 := (2\pi)^{m/2} e^{-\frac{t^2}{4}}$, then

$$M^\alpha_{\mathfrak{a}_0} = \int_{\mathbb{R}^m} \xi^\gamma e^{-|\xi|^2/2} d\xi = (2\pi)^{m/2} \prod_{j=1}^m \int_{\mathbb{R}} \xi^{\alpha_j} \boldsymbol{\gamma}\big[\,d\xi\,\big],$$

where $\boldsymbol{\gamma}$ denotes the Gaussian measure on $\mathbb{R}$ with mean zero and variance 1. If α is even, $\alpha = 2\kappa$, then

$$M^{\mathfrak{a}_0}_{2\kappa} \stackrel{(1.1.9)}{=} (2\pi)^{m/2} \prod_{j=1}^m (2\kappa_j - 1)!!.$$

On the other hand, using (2.3.12) we deduce

$$\begin{aligned}M^{\mathfrak{a}_0}_{2\kappa} &= \boldsymbol{m}_{2\kappa} \int_0^\infty r^{m+2|\kappa|-1} e^{-r^2/2} dr = \sqrt{\frac{\pi}{2}} \boldsymbol{m}_{2\kappa} \int_{\mathbb{R}} |x|^{m+2|\kappa|-1} \boldsymbol{\gamma}\big[\,dx\,\big] \\ &\stackrel{(1.1.6)}{=} 2^{|\kappa|+m/2-1} \boldsymbol{m}_{2\kappa} \Gamma\big(\,|\kappa| + m/2\,\big).\end{aligned}$$

Hence

$$\boldsymbol{m}_{2\kappa} = 2^{|\kappa|+m/2-1} \frac{\prod_{j=1}^m (2\kappa_j - 1)!!}{\Gamma\big(\,|\kappa| + m/2\,\big)} = \frac{2 \prod_{j=1} \Gamma(\kappa_j + 1/2)}{\Gamma\big(\,|\kappa| + m/2\,\big)}. \tag{2.3.13}$$

For every $k \in \mathbb{Z}_{\geq 0}$ we set

$$I_k(\mathfrak{a}) := \int_0^\infty r^k \mathfrak{a}(r)^2 dr.$$

We deduce

$$(2\pi)^m M_{\mathfrak{a}}^{2\kappa} = 2I_{m-1+2|\kappa|}(\mathfrak{a}) \frac{2\prod_{j=1}\Gamma(\kappa_j+1/2)}{\Gamma\big(\,|\kappa|+m/2\,\big)}. \tag{2.3.14}$$

We set

$$s_m := \int_{\mathbb{R}^m} \mu_{\mathfrak{a}}\big[\,d\xi\,\big],\;\; d_m := \int_{\mathbb{R}^m} \xi_1^2\mu_{\mathfrak{a}}\big[\,d\xi\,\big],\;\; h_m := \int_{\mathbb{R}^m} \xi_1^2\xi_2^2\mu_{\mathfrak{a}}\big[\,d\xi\,\big]. \tag{2.3.15}$$

Then

$$\int_{\mathbb{R}^m} \mathfrak{a}(|\xi|)^2 d\xi = \frac{2\pi^{m/2}}{\Gamma(m/2)} I_{m-1}(\mathfrak{a}) = (2\pi)^m s_m, \tag{2.3.16}$$

$$\int_{\mathbb{R}^m} \xi_j^2\mathfrak{a}(|\xi|)^2 d\xi = \frac{2\pi^{m/2}}{\Gamma(m/2+1)} I_{m+1}(\mathfrak{a}) = (2\pi)^m d_m,\;\; \forall j, \tag{2.3.17}$$

$$\int_{\mathbb{R}^m} \xi_j^2\xi_k^2\mathfrak{a}(|\xi|)^2 d\xi = \frac{(2\pi)^{m/2}}{\Gamma(m/2+2)} I_{m+3}(\mathfrak{a}) = (2\pi)^m h_m,\;\; \forall j\neq k, \tag{2.3.18}$$

$$\int_{\mathbb{R}^m} \xi_j^4\mathfrak{a}(|\xi|)^2 d\xi = \frac{6\pi^{m/2}}{\Gamma(m/2+1)} I_{m+3}(\mathfrak{a}) = 3(2\pi)^m h_m,\;\; \forall j. \tag{2.3.19}$$

Using (2.3.10) and (2.3.11) we deduce that for any $\boldsymbol{x}\in\mathbb{R}^m$ *the Gaussian vectors* $\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})$ *and* $\operatorname{Hess}_{\Phi_{\mathfrak{a}}}(\boldsymbol{x})$ *are independent*. Hence, the Kac-Rice density

$$\rho_{\mathfrak{a}}(\boldsymbol{x}) = \mathbb{E}\big[\,|\det \operatorname{Hess}_{\Phi_{\mathfrak{a}}}(\boldsymbol{x})|\;\big|\;\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})=0\,\big]p_{\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})}(0)$$

simplifies to

$$\rho_{\mathfrak{a}}(\boldsymbol{x}) = \mathbb{E}\big[\,|\det \operatorname{Hess}_{\Phi_{\mathfrak{a}}}(\boldsymbol{x})|\,\big]p_{\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})}(0).$$

Using (2.3.10)and (2.3.17) we deduce that the variance operator of $\nabla\Phi(\boldsymbol{x})$ is

$$\operatorname{Var}\big[\,\nabla\Phi(\boldsymbol{x})\,\big] = d_m\mathbb{1}_m,\;\; \forall \boldsymbol{x}\in\mathbb{R}^m. \tag{2.3.20}$$

In particular, this show that $\nabla\Phi_{\mathfrak{a}}$ is an ample random field and thus $\Phi_{\mathfrak{a}}$ is a.s. Morse.

As explained in Example 2.3.6, the space $\mathbf{Sym}(\mathbb{R}^m)$ of real symmetric $m\times m$ matrices is equipped with an inner product $(A,B)=\operatorname{tr}(AB)$. The linear functions

$$\ell_{ij},\omega_{ij} : \mathbf{Sym}(\mathbb{R}^m)\to\mathbb{R},\;\; 1\leq i\leq j\leq m,$$

$$\ell_{ij}(A) = a_{ij},\;\; \omega_{ij}(A) = \begin{cases} a_{ii}, & i=j,\\ \sqrt{2}a_{ij}, & i<j \end{cases} \tag{2.3.21}$$

define coordinates on $\mathbf{Sym}(\mathbb{R}^m)$. Additionally $\big(\omega_{ij}\big)_{1\leq i\leq j\leq m}$ are *orthonormal* with respect to the above inner product. We set

$$L_{ij}(\boldsymbol{x}) := \ell_{ij}\big(\operatorname{Hess}_{\Phi_{\mathfrak{a}}}(\boldsymbol{x})\big),\;\; \Omega_{ij}(\boldsymbol{x}) := \omega_{ij}\big(\operatorname{Hess}_{\Phi_{\mathfrak{a}}}(\boldsymbol{x})\big). \tag{2.3.22}$$

Then

$$\mathbb{E}\big[\,\partial^2_{x_ix_j}\Phi_{\mathfrak{a}}(x)\partial^2_{x_kx_\ell}\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m}\xi_i\xi_j\xi_k\xi_\ell a\big(\,|\xi|^2\,\big)d\xi,\;\; i\leq j,\;\; k\leq\ell.$$

Note that if $i<j$, then the above integral is nonzero iff $(i,j)=(k,\ell)$ in which case

$$\begin{aligned}\mathbb{E}\big[\,L_{ij}(\boldsymbol{x})L_{k\ell}(\boldsymbol{x})\,\big] &= \mathbb{E}\big[\,\big(\,\partial^2_{x_ix_j}\Phi_{\mathfrak{a}}(x)\,\big)^2\,\big]\\ &= \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m}\xi_i^2\xi_j^2\mathfrak{a}\big(\,|\xi|^2\,\big)d\xi \overset{(2.3.18)}{=} h_m.\end{aligned}$$

If $i=j$ then the above integral is nonzero iff $k=\ell$, in which case we deduce from (2.3.18) and (2.3.19)

$$\mathbb{E}\big[\,\partial^2_{x_i}\Phi_{\mathfrak{a}}(x)\partial^2_{x_k}\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big]=\begin{cases}h_m & i\neq k,\\ 3h_m, & i=k.\end{cases}$$

The above equalities can be rewritten in the more compact form

$$\mathbb{E}\big[\,L_{ij}(\boldsymbol{x})L_{k\ell}(\boldsymbol{x})\,\big]=h_m\big(\,\delta_{ij}\delta_{k\ell}+\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk}\,\big),\ \ \forall i\leq j,\ k\leq\ell. \tag{2.3.23}$$

These equalities show that the off diagonal entries of $\mathrm{Hess}_{\Phi_{\mathfrak{a}}}$ are i.i.d., and also independent of the diagonal entries. The diagonal entries have identical distributions but are dependent. The parameter h_m describes the various variances and covariances.

The Gaussian measure on $\mathbf{Sym}(\mathbb{R}^m)$ determined by these covariance equalities is invariant with respect to the action by conjugation of $O(m)$ on $\mathbf{Sym}(\mathbb{R}^m)$. This corresponds to the Gaussian ensemble of random matrices denoted by $\mathcal{S}_m^{h_m,h_m}$ in Appendix C.1. For ease of notation we set $\mathcal{S}_m^{h_m}:=\mathcal{S}_m^{h_m,h_m}$. In particular this proves that

$$\forall\boldsymbol{x}\in\mathbb{R}^m,\ \ \rho_{\mathfrak{a}}(\boldsymbol{x})=\rho_{\mathfrak{a}}(0)=(2\pi d_m)^{-m/2}\mathbb{E}_{\mathcal{S}_m^{h_m}}\big[\,|\det H|\,\big] \tag{2.3.24}$$

We deduce from the Kac-Rice formula (**KR**) that for any box B

$$\mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[B]\,\big]=\int_B\rho_{\mathfrak{a}}(\boldsymbol{x})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}\,\big]=\rho_{\mathfrak{a}}(0)\boldsymbol{\lambda}\big[\,B\,\big]=(2\pi d_m)^{-m/2}\mathbb{E}_{\mathcal{S}_m^{h_m}}\big[\,|\det H|\,\big]\boldsymbol{\lambda}\big[\,B\,\big]$$

$(X=(2h_m)^{-1/2}H)$

$$=\left(\frac{h_m}{\pi d_m}\right)^{m/2}\mathbb{E}_{\mathcal{S}_m^{1/2}}\big[\,|\det X|\,\big]\boldsymbol{\lambda}\big[\,B\,\big].$$

Lemma C.1.2 with $u=v=\frac{1}{2}$ implies that

$$\mathbb{E}_{\mathcal{S}_m^{1/2}}\big[\,|\det H|\,\big]=2^{\frac{3}{2}}\Gamma\left(\frac{m+3}{2}\right)\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}}\rho_{m+1,1/2}(x)e^{-\frac{x^2}{2}}dx,$$

where $\rho_{N,v}(\lambda)$ denotes the normalized 1-point correlation function of the Gaussian Orthogonal Ensemble $GOE_N^v=\mathcal{S}_N^{0,v}$; see Appendix C.1.

Using the equality (C.1.10) we deduce that

$$n^{1/2}\rho_{n,1/2}\big(\,n^{1/2}y\,\big)=\rho_{n,1/2n}(y),$$

so that

$$\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}}\rho_{m+1,1/2}(x)e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}}\rho_{m+1,\frac{1}{2(m+1)}}(y)e^{-\frac{(m+1)y^2}{2}}dy$$

$$=\left(\frac{2}{m+1}\right)^{1/2}\int_{\mathbb{R}}\rho_{m+1,\frac{1}{2(m+1)}}(y)\boldsymbol{\gamma}_{\frac{1}{(m+1)}}\big[\,dy\,\big].$$

We deduce form (C.1.15) that

$$\int_{\mathbb{R}}\rho_{m+1,\frac{1}{2(m+1)}}(y)\boldsymbol{\gamma}_{\frac{1}{2(m+1)}}\big[\,dy\,\big]\sim\frac{\sqrt{2}}{\pi}\ \text{ as } m\to\infty.$$

Hence

$$\mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[B]\,\big]=C_m(\mathfrak{a})\boldsymbol{\lambda}\big[\,B\,\big] \tag{2.3.25}$$

where

$$\begin{aligned} C_m(\mathfrak{a}) &= \rho_{\mathfrak{a}}(0) = (2\pi d_m)^{-m/2}\mathbb{E}_{\mathcal{S}_m^{h_m}}\big[\,|\det H|\,\big] \\ &= 2^{\frac{3}{2}}\Gamma\left(\frac{m+3}{2}\right)\left(\frac{h_m}{\pi d_m}\right)^{m/2}\left(\frac{2}{m+1}\right)^{1/2}\int_{\mathbb{R}}\rho_{m+1,\frac{1}{2(m+1)}}(y)\gamma_{\frac{1}{(m+1)}}\big[dy\big]. \\ &\sim 2^{\frac{5}{2}}\Gamma\left(\frac{m+3}{2}\right)\left(\frac{h_m}{\pi d_m}\right)^{m/2}\left(\frac{1}{m+1}\right)^{1/2} \quad \text{as } m\to\infty \end{aligned} \tag{2.3.26}$$

Using (2.3.17) and (2.3.18) we deduce

$$\frac{h_m}{d_m} = \frac{\Gamma(1+m/2)}{\Gamma(2+m/2)}\times\frac{I_{m+3}(\mathfrak{a})}{I_{m+1}(\mathfrak{a})} = \frac{2I_{m+3}(\mathfrak{a})}{(m+2)I_{m+1}(\mathfrak{a})}.$$

Hence

$$\begin{aligned} C_m(\mathfrak{a}) &\sim 2^{5/2}\left(\frac{h_m(\mathfrak{a})}{d_m(\mathfrak{a})}\right)^{m/2}\Gamma\left(\frac{m+3}{2}\right)m^{-1/2} \\ &\sim 2^{\frac{m+5}{2}}\left(\frac{I_{m+3}(\mathfrak{a})}{(m+2)I_{m+1}(\mathfrak{a})}\right)^{m/2}\Gamma\left(\frac{m+3}{2}\right)m^{-1/2} \quad \text{as } m\to\infty. \end{aligned} \tag{2.3.27}$$

The constant $C_m(\mathfrak{a})$ tends to grow very fast as $m\to\infty$, but its large m behavior depends on the tail of the amplitude $\mathfrak{a}$. Roughly speaking, the slower the decay at ∞ of $\mathfrak{a}$ the faster the growth of $C_m(\mathfrak{a})$. Here are some examples. Appendix B.4 contains their proofs.

- If $\mathfrak{a}(t)^2 = e^{-t^2}$, then
$$\log C_m(\mathfrak{a}) \sim \frac{m}{2}\log m \;\; \text{as } m\to\infty.$$
- If
$$\mathfrak{a}(t)^2 = \exp\big(-(\log t)\log(\log t)\,\big),\;\; \forall t\geq 1,$$
then
$$\log C_m(\mathfrak{a}) \sim \frac{m}{2}e^{m+2}(e^2-1),\;\; \text{as } m\to\infty.$$
- Fix $C>0$ and $\alpha>1$. If
$$\mathfrak{a}(t)^2 = \exp\big(-C(\log t)^{\alpha}\,\big),\;\; \forall t>1,$$
then
$$\log C_m(\mathfrak{a}) \sim \frac{Z(\alpha,C)}{\alpha-1}m^{\frac{\alpha}{\alpha-1}},\;\; \text{as } m\to\infty,$$
where $Z(\alpha,C)$ is a positive constant depending explicitly on α and C.

□

Example 2.3.8 (Random Fourier series)**.** Fix an amplitude $\mathfrak{a}$, i.e., and even Schwartz function $\mathfrak{a}\in\mathcal{S}(\mathbb{R})$ such $\mathfrak{a}(0)=1$ and consider the Gaussian function $F^R_{\mathfrak{a}}(\vec{\theta})$ defined in in Example 1.2.31. More precisely,

$$F^R_{\mathfrak{a}}(\vec{\theta}) = R^{-m/2}\Big(\,A_0u_0 + \sum_{\vec{\ell}\succ 0}\mathfrak{a}\big(\,|\,2\pi\vec{\ell}\,|/R\big)\big(\,A_{\vec{\ell}}u_{\vec{\ell}}(\vec{\theta}) + B_{\vec{\ell}}v_{\vec{\ell}}(\vec{\theta})\,\big)\,\Big),$$

where $A_{\vec{\ell}}$, $B_{\vec{k}}$, $\vec{\ell}\succeq 0$, $\vec{k}\succ 0$ are independent standard normal random variables. We view $F^R_{\mathfrak{a}}$ as a $\mathbb{Z}^m$-periodic random smooth function on $\mathbb{R}^m$.

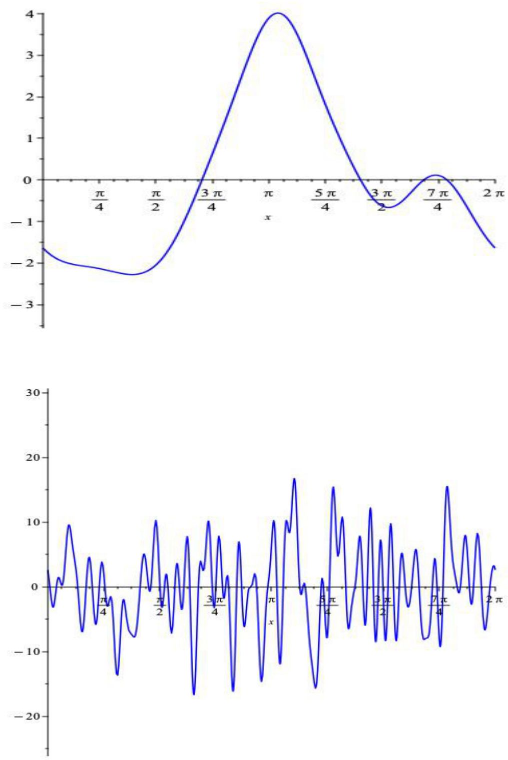

Figure 2.2. *Samples of $F^R_{\mathfrak{a}}$ for $m = 1$, $\mathfrak{a}(t) = e^{-t^2}$. At the top $R = 10$ and at the bottom $R = 100$.*

As in the previous example we set

$$w_{\mathfrak{a}}(\xi) = w_{\mathfrak{a},m}(\xi) := \mathfrak{a}\big(\,|\,\xi\,|\,\big)^2,$$

and define as in (2.3.8)

$$\boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{x}\,\big) = \frac{1}{(2\pi)^m}\widehat{w}_{\mathfrak{a}}(\boldsymbol{x}) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle} w_{\mathfrak{a},m}(\xi)d\xi.$$

In (1.2.25) we showed that the covariance kernel $\mathcal{C}^R_{\mathfrak{a}}$ of $F^R_{\mathfrak{a}}$ admits the description

$$\begin{aligned}\mathcal{C}^R_{\mathfrak{a}}(\vec{\varphi}+\vec{\tau},\vec{\varphi}) = C^R_{\mathfrak{a}}(\vec{\tau}) &= \frac{1}{(2\pi)^m}\sum_{\vec{k}\in\mathbb{Z}^m}\widehat{w}_{\mathfrak{a}}\big(\,(\vec{k}-\vec{\tau})R\,\big)\\ &= \boldsymbol{K}_{\mathfrak{a}}\big(\,\vec{\tau}R\,\big) + \underbrace{\sum_{\vec{k}\in\mathbb{Z}^m\setminus 0}\boldsymbol{K}_{\mathfrak{a}}\big(\,(\vec{k}-\vec{\tau})R\big)}_{=:\mathcal{E}_R(\vec{\tau})}.\end{aligned}$$

I want to investigate the distribution $\mathfrak{C}\big[\,-, F^R_{\mathfrak{a}}\big]$ of critical points of $F^R_{\mathfrak{a}}$.

The random function $F^R_{\mathfrak{a}}$ is highly oscillatory as $R\nearrow\infty$ so we expect that it will have more and more critical points as R increases; see Figure 2.2.

Before we proceed further we need to introduce some terminology. We will denote by $O\big(R^{-\infty}\big)$ any quantity $q(R)$ such that, for any $N\in\mathbb{N}$, we have

$$q\big(R\big) = O\big(R^{-N}\big) \ \text{ as } R\to\infty.$$

Fix a box $B\subset[0,1]^m\subset\mathbb{R}^n$. Since $\boldsymbol{K}_{\mathfrak{a}}$ is a Schwartz function we deduce that for any multi-index $\alpha\in\big(\mathbb{Z}_{\geq 0}\big)^m$

$$\partial^\alpha_{\vec{\tau}} K^R_{\mathfrak{a}}(0) = R^{|\alpha|}\big(\boldsymbol{K}_{\mathfrak{a}}(0)+O(R^{-\infty})\big), \ \text{ as } \hbar\searrow 0. \tag{2.3.28}$$

We have

$$\mathbb{E}\big[\,\partial^\alpha_{\vec{\theta}} F^R_{\mathfrak{a}}(\vec{\theta})\partial^\beta_{\vec{\varphi}} F^R_{\mathfrak{a}}(\vec{\varphi})\,\big] = \partial^\alpha_{\vec{\theta}}\partial^\beta_{\vec{\varphi}}\mathfrak{C}^R_{\mathfrak{a}}\big(\vec{\theta},\vec{\varphi}\big),$$

and thus

$$\mathbb{E}\big[\,\partial^\alpha F^R_{\mathfrak{a}}(\vec{\theta})\partial^\beta F^R_{\mathfrak{a}}(\vec{\varphi})\,\big]_{\vec{\theta}=\vec{\varphi}} = \partial^\alpha_{\vec{\theta}}\partial^\beta_{\vec{\varphi}}\mathfrak{C}^R_{\mathfrak{a}}\big(\vec{\theta},\vec{\varphi}\big)_{\vec{\theta}=\vec{\varphi}} = (-1)^{|\beta|}\partial^{\alpha+\beta}_{\vec{\tau}} C^R_{\mathfrak{a}}(0).$$

Observe that since $C^R_{\mathfrak{a}}(\vec{\tau}) = C^R_{\mathfrak{a}}(-\vec{\tau})$ we have

$$\partial^\alpha_{\vec{\theta}}\partial^\beta_{\vec{\varphi}}\mathfrak{C}^R_{\mathfrak{a}}\big(\vec{\theta},\vec{\varphi}\big)_{\vec{\theta}=\vec{\varphi}} = 0$$

if $|\alpha|+|\beta|$ is odd. Hence

$$\nabla F^R_{\mathfrak{a}}\big(\vec{\theta}\big) \text{ and } \operatorname{Hess}_{F^R_{\mathfrak{a}}}\big(\vec{\theta}\big) \text{ are independent Gaussian vectors for any } \vec{\theta}. \tag{2.3.29}$$

We deduce from (2.3.28) that

$$(-1)^{|\beta|}\partial^{\alpha+\beta}_{\vec{\tau}} C^R_{\mathfrak{a}}(0) = R^{|\alpha|+|\beta|}\Big(\,(-1)^{|\beta|}\partial^{\alpha+\beta}_{\vec{\tau}}\boldsymbol{K}_{\mathfrak{a}}(0) + O\big(R^{-\infty}\big)\Big).$$

On the other hand,

$$(-1)^{|\beta|}\partial^{\alpha+\beta}_{\vec{\tau}}\boldsymbol{K}_{\mathfrak{a}}(0) = \mathbb{E}\big[\,\partial^\alpha\Phi_{\mathfrak{a}}(\vec{\theta})\partial^\beta\Phi_{\mathfrak{a}}(\varphi)\,\big]_{\vec{\theta}=\varphi}.$$

Hence

$$\operatorname{Var}\big[\,R^{-1}\nabla F^R_{\mathfrak{a}}\,\big] = \operatorname{Var}\big[\,\nabla\Phi_{\mathfrak{a}}\,\big] + O\big(R^{-\infty}\big)$$

and

$$\operatorname{Var}\big[\,R^{-2}\operatorname{Hess}_{F^R_{\mathfrak{a}}}\,\big] = \operatorname{Var}\big[\,\operatorname{Hess}_{\Phi_{\mathfrak{a}}}\,\big] + O\big(R^{-\infty}\big).$$

We have computed the covariances $\operatorname{Var}\big[\,\nabla\Phi_{\mathfrak{a}}\,\big]$ $\operatorname{Var}\big[\,\operatorname{Hess}_{\Phi_{\mathfrak{a}}}\,\big]$ in Example 2.3.7.

We set

$$\mathfrak{C}^R_{\mathfrak{a}} := \mathfrak{C}[-,F^R_{\mathfrak{a}}].$$

Fix a box $B\subset[0,1]^m$. Since $F^R_{\mathfrak{a}}$ is stationary, the Kac-Rice density $\rho^{\nabla F^R_{\mathfrak{a}}}_{KR}$ is constant and we deduce

$$\mathbb{E}\big[\,\mathfrak{C}^R_{\mathfrak{a}}[B]\,\big] = \mathbb{E}\big[\,|\det\operatorname{Hess}_{F^R_{\mathfrak{a}}}(0)|\,\big]p_{\nabla F^R_{\mathfrak{a}}(0)}(0)\boldsymbol{\lambda}\big[\,B\,\big].$$

We have

$$\mathbb{E}\big[\,|\det\operatorname{Hess}_{F^R_{\mathfrak{a}}}(0)|\,\big] = R^{2m}\mathbb{E}\big[\,|\det R^{-2}\operatorname{Hess}_{F^\hbar_{\mathfrak{a}}}(0)|\,\big],$$

$$p_{\nabla F^R_{\mathfrak{a}}(0)}(0) = R^m p_{R^{-1}\nabla F^R_{\mathfrak{a}}(0)}(0) + O\big(R^{-\infty}\big).$$

Hence

$$\mathbb{E}\big[\,\mathfrak{C}^R_{\mathfrak{a}}[B]\,\big] = R^{2m}\mathbb{E}\big[\,|\det\operatorname{Hess}_{F^R_{\mathfrak{a}}}(0)|\,\big]p_{R^{-1}\nabla F^R_{\mathfrak{a}}(0)}(0)\boldsymbol{\lambda}\big[\,B\,\big]$$

Since

$$\mathbb{E}\big[\,|\det R^{-2}\operatorname{Hess}_{F^R_{\mathfrak{a}}}(0)|\,\big]p_{\hbar\nabla F^R_{\mathfrak{a}}(0)}(0) = \mathbb{E}\big[\,|\det\operatorname{Hess}_{\Phi_{\mathfrak{a}}}(0)|\,\big]p_{\nabla\Phi_{\mathfrak{a}}}(0) + O\big(R^{-\infty}\big)$$

we deduce

$$\begin{aligned}\mathbb{E}\big[\,\mathfrak{C}^R_{\mathfrak{a}}[B]\,\big] &= R^m\big(\,\mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[B]\,\big] + O\big(\,R^{-\infty}\,\big)\\ &= R^m C_m(\mathfrak{a})\operatorname{vol}\big[\,B\,\big] + O\big(\,R^{-\infty}\,\big).\end{aligned} \tag{2.3.30}$$

More generally, if $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$ we have

$$\begin{aligned}\mathbb{E}\big[\,\mathfrak{C}^R_{\mathfrak{a}}[f]\,\big] &= \mathbb{E}\big[\,|\det \operatorname{Hess}_{F^R_{\mathfrak{a}}}(0)|\,\big] p_{\nabla F^R_{\mathfrak{a}}(0)}(0)\int_{\mathbb{R}^m} f(\boldsymbol{x})d\boldsymbol{x}\\ &= R^m\Big(\,C_m(\mathfrak{a}) + O\big(\,R^{-\infty}\,\big)\,\Big)\int_{\mathbb{R}^m} f(\boldsymbol{x})d\boldsymbol{x}.\end{aligned} \tag{2.3.31}$$

□

2.3.3. The distribution of critical values. Consider the isotropic Gaussian function $\Phi_{\mathfrak{a}} : \mathbb{R}^m \to \mathbb{R}$ whose critical points were investigated in Example 2.3.7.

We recall that $\mathfrak{a} \in \mathcal{S}\big(\,\mathbb{R}\,\big)$ is an even Schwartz function such that $\mathfrak{a}(0) = 1$. The function $\Phi_{\mathfrak{a}}$ is determined by the covariance kernel

$$\mathcal{K}^{\mathfrak{a}}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{K}_{\alpha}\big(\,\boldsymbol{x}-\boldsymbol{y}\,\big),\;\; \boldsymbol{K}_{\mathfrak{a}}\big(\,\boldsymbol{x}\,\big) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mu_{\mathfrak{a}}\big[\,d\xi\,\big], \tag{2.3.32}$$

where the spectral measure $\mu_{\mathfrak{a}} \in \operatorname{Meas}\big(\,\mathbb{R}^m\,\big)$ is given by

$$\mu_{\mathfrak{a}}\big[\,d\xi\,\big] = \frac{1}{(2\pi)^m} w_{\mathfrak{a}}(\xi)\boldsymbol{\lambda}\big[\,d\xi\,\big],\;\; w_{\mathfrak{a}}(\xi) = w_{\mathfrak{a},m}(\xi) = \mathfrak{a}\big(\,|\xi|\,\big)^2.$$

The random function $\Phi_{\mathfrak{a}}$ is a.s. Morse. We set $\mathfrak{C}_{\mathfrak{a}} = \mathfrak{C}[-,\Phi_{\mathfrak{a}}]$. Thus, for each box $B \subset \mathbb{R}^m$, $\mathfrak{C}_{\mathfrak{a}}[B]$ is the number of critical points of $\Phi_{\mathfrak{a}}$ inside B. According to (2.3.25) we have

$$\mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[B]\,\big] = C_m(\mathfrak{a})\boldsymbol{\lambda}\big[\,B\,\big],$$

where $C_m(\mathfrak{a})$ is the positive constant described explicitly in (2.3.26) and $\boldsymbol{\lambda}\big[\,B\,\big]$ is the (Lebesgue) volume of B. In particular, $\mathfrak{C}_{\mathfrak{a}}[B]$ is a.s. finite. We denote by $\mathfrak{D}_{B,\mathfrak{a}}$ the random measure on the real axis

$$\mathfrak{D}_{B,\mathfrak{a}} := \sum_{\boldsymbol{x}\in B\cap\nabla^{\Phi}_{\mathfrak{a}}(0)} \delta_{\Phi_{\mathfrak{a}}(\boldsymbol{x})} \in \operatorname{Meas}\big(\,\mathbb{R}\,\big)$$

I will refer to $\mathfrak{D}_{B,\mathfrak{a}}$ as the *discriminant measure* of $\Phi_{\mathfrak{a}}|_B$. It is supported on the set of critical values of $\Phi_{\mathfrak{a}}|_B$. In singularity theory this set is usually referred to as the *discriminant locus*[4] of $\Phi_{\mathfrak{a}}$.

For every Borel set $C \subset \mathbb{R}$ we have

$$\mathfrak{D}_{B,\mathfrak{a}}\big[\,C\,\big] = \sum_{\boldsymbol{x}\in B\cap\nabla\Phi_{\mathfrak{a}}(0)} \boldsymbol{I}_C\big(\,\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big) = \#\big\{\,\boldsymbol{x}\in B;\;\; \nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}) = 0,\;\; \Phi_{\mathfrak{a}}(\boldsymbol{x}) \in C\,\big\}.$$

Note

$$\mathfrak{D}_{B,\mathfrak{a}}\big[\,\mathbb{R}\,\big] = \mathfrak{C}_{\mathfrak{a}}\big[\,B\,\big].$$

This subsection is devoted to an investigation of the random measure $\mathfrak{D}_{\mathfrak{a},B}$. Note that the expectation

$$\nu_{\mathfrak{a}}\big[\,C\,\big] := \mathbb{E}\Big[\,\mathfrak{D}_{B,\mathfrak{a}}\big[\,C\,\big]\,\Big] \tag{2.3.33}$$

[4]The term "locus" is meant to suggest that this set has additional structure. In algebraic geometry it is a a scheme. In our looser context the additional structure is a measure supported on this set.

is a Borel measure on $\mathbb{R}$. It describes the expected number of the critical values of $\Phi_{\mathfrak{a}}|_B$ that are located in C. Set

$$\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x}) := \operatorname{Hess}_{\Phi_{\mathfrak{a}}}(\boldsymbol{x}),\ \ \boldsymbol{x} \in \mathbb{R}^m.$$

In Example 1.2.35 we proved that $\Phi_{\mathfrak{a}}$ is J_1-ample so the Gaussian vector

$$W(\boldsymbol{x}) = W_{\mathfrak{a}}(\boldsymbol{x}) := \Phi_{\mathfrak{a}}(\boldsymbol{x}) \oplus \nabla\Phi_{\mathfrak{a}}(z) \tag{2.3.34}$$

is nondegenerate for any $\boldsymbol{x} \in \mathbb{R}^m$.

The Gaussian random variable $\Phi_{\mathfrak{a}}(\boldsymbol{x})$ has variance

$$\mathbb{E}\big[\,\Phi_{\mathfrak{a}}(0)^2\,\big] = \boldsymbol{K}_{\mathfrak{a}}(0) \stackrel{(2.3.15)}{=} s_m.$$

Hence

$$\operatorname{Var}\big[\,W(\boldsymbol{x})\,\big] = \begin{bmatrix} s_m & 0 \\ 0 & d_m\mathbb{1}_m \end{bmatrix}$$

We can thus invoke (2.2.11b) to deduce that if $C \subset \mathbb{R}$ is a compact interval, then

$$\nu_{\mathfrak{a}}\big[\,C\,\big] = \int_B r(\boldsymbol{x})(\boldsymbol{x})p_{\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})}(0)d\boldsymbol{x},$$

where

$$r(\boldsymbol{x}) = \int_C \mathbb{E}\big[\,|\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})|\,\big|\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}) = 0,\ \Phi_{\mathfrak{a}}(\boldsymbol{x}) = t\,\big]\mathbb{P}_{\Phi_{\mathfrak{a}}(\boldsymbol{x})}\big[\,dt\,\big].$$

As shown above $\mathbb{P}_{\Phi_{\mathfrak{a}}(\boldsymbol{x})} = \boldsymbol{\Gamma}_{s_m}$, $\forall\boldsymbol{x}$. We deduce from Fubini's theorem that

$$\nu_{\mathfrak{a}}\big[\,C\,\big] = \int_C \rho_{\mathfrak{a}}(t)\boldsymbol{\Gamma}_{s_m}\big[\,dt\,\big], \tag{2.3.35}$$

where

$$\begin{aligned} \rho_{\mathfrak{a}}(t) &= \int_B \mathbb{E}\big[\,|\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})|\,\big|\Phi_{\mathfrak{a}}(\boldsymbol{x}) = t,\ \nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}) = 0\,\big]p_{\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})}(0)d\boldsymbol{x} \\ &\stackrel{(2.3.20)}{=} \big(\,2\pi d_m\,\big)^{-m/2}\int_B \mathbb{E}\big[\,|\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})|\,\big|W(\boldsymbol{x}) = (t,0)\,\big]d\boldsymbol{x} \\ &= \big(\,2\pi d_m\,\big)^{-m/2}\mathbb{E}\big[\,|\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})|\,\big|W(\boldsymbol{x}) = (t,0)\,\big]\operatorname{vol}\big[\,B\,\big]. \end{aligned} \tag{2.3.36}$$

At the last step I have used the stationarity of $\Phi_{\mathfrak{a}}$ that implies that the integrand in the second equality is independent of $\boldsymbol{x}$. To compute the above conditional expectation I will rely on the Gaussian regression formula.

The variance of $\operatorname{Hess}_{\mathfrak{a}}$ is given by (2.3.23)

$$\mathbb{E}\big[\,L_{ij}(\boldsymbol{x})L_{k\ell}(\boldsymbol{x})\,\big] = h_m\big(\,\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}\,\big),\ \ \forall i \leq j,\ k \leq \ell$$

where $L_{ij}(bx)$, and $\Omega_{ij}(\boldsymbol{x})$ are defined by (2.3.22). Set

$$W = (W_0, W_1, \dots, W_m),\ \ W_0 = \Phi_{\mathfrak{a}}(\boldsymbol{x}),\ \ W_j = \partial_{x_j}\Phi_{\mathfrak{a}}(\boldsymbol{x}).$$

Denote by $\overline{\operatorname{Hess}}_{\mathfrak{a}}(\boldsymbol{x})$ the random symmetric matrix with variance given by the regression formula

$$\operatorname{Var}\big[\overline{\operatorname{Hess}}_{\mathfrak{a}}(\boldsymbol{x})\,\big] = \operatorname{Var}\big[\,\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})\,\big] - \operatorname{Cov}\big[\,\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x}), Y\,\big]\operatorname{Var}\big[\,W\,\big]^{-1}\operatorname{Cov}\big[\,W, \operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})\,\big].$$

Set

$$\bar{L}_{ij} = \ell_{ij}\big(\overline{\operatorname{Hess}}_{\mathfrak{a}}(\boldsymbol{x})\,\big),\ \ \bar{\Omega}_{ij} := \omega_{ij}\big(\overline{\operatorname{Hess}}_{\mathfrak{a}}(\boldsymbol{x})\,\big),$$

$$C_{ij|a} := \operatorname{Cov}\big[\,\Omega_{ij}, W_a\,\big],\ \ 1 \leq i \leq j \leq m,\ \ 0 \leq a \leq m.$$

If we write

$$\operatorname{Var}\big[\,W\,\big]^{-1}=\big(\,t_{ab}\,\big)_{0\leq a,b\leq m},$$

then

$$\mathbb{E}\big[\,\bar{\Omega}_{ij}\bar{\Omega}_{k\ell}\,\big]=\mathbb{E}\big[\,\Omega_{ij}(\boldsymbol{x})\Omega_{k\ell}(\boldsymbol{x})\,\big]-\sum_{a,b=0}^{m}C_{ij|a}t_{ab}C_{k\ell|b}. \tag{2.3.37}$$

Since $\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})$ and $\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})$ are independent we deduce

$$C_{ij|a}=0,\;\;\forall a=1,\dots,m.$$

Hence

$$\mathbb{E}\big[\,\bar{\Xi}_{ij}\bar{\Xi}_{k\ell}\,\big]=\mathbb{E}\big[\,\Xi_{ij}(\boldsymbol{x})\Xi_{k\ell}(\boldsymbol{x})\,\big]-\frac{1}{s_m}C_{ij|0}C_{k\ell|0}. \tag{2.3.38}$$

Observe that if $i\neq j$, then

$$C_{ij|0}=\frac{-\sqrt{2}}{(2\pi)^m}\int_{\mathbb{R}^m}\xi_i\xi_j\mathfrak{a}\big(\,|\xi|^2\,\big)d\xi=0.$$

Moreover

$$C_{jj|0}=-\frac{1}{(2\pi)^m}\int_{\mathbb{R}^m}\xi_j^2\mathfrak{a}\big(\,|\xi|^2\,\big)d\xi=-d_m.$$

We deduce from (2.3.37) and (2.3.38) that

$$\mathbb{E}\big[\,\bar{L}_{ij}\bar{L}_{k\ell}\,\big]=\underbrace{\left(h_m-\frac{d_m^2}{s_m}\right)}_{=:u_m}\delta_{ij}\delta_{k\ell}+h_m\big(\,\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk}\,\big),\;\;\forall i\leq j,\;\;k\leq\ell. \tag{2.3.39}$$

These equalities determine a $O(m)$-invariant Gaussian measure $\boldsymbol{\Gamma}_{u_m,h_m}$ on $\mathbf{Sym}(\mathbb{R}^m)$; see (C.1.3). In Appendix C.1 the probability space $\big(\,\mathbf{Sym}(\mathbb{R}^m),\boldsymbol{\Gamma}_{u_m,h_m}\,\big)$ obtained in this fashion is denoted by $\mathcal{S}_m^{u_m,h_m}$.

To apply the Gaussian regression formula (1.1.20) we need to use the regression operator

$$R_{\operatorname{Hess}_a,W}=\operatorname{Cov}\big[\,\operatorname{Hess}_a,W\,\big]\operatorname{Var}\big[\,W\,\big]^{-1}:\mathbb{R}^{m+1}\to\mathbf{Sym}(\mathbb{R}^m).$$

We have

$$R_{\operatorname{Hess}_a,W}\begin{bmatrix}t\\0\\\vdots\\0\end{bmatrix}=C_{11|0}s_m^{-1}t\mathbb{1}_m=-\frac{d_mt}{s_m}\mathbb{1}_m$$

Using the regression formula we deduce

$$\mathbb{E}\big[\,|\operatorname{Hess}_{\mathfrak{a}}(\boldsymbol{x})|\,\big|\,W(\boldsymbol{x})=(t,0)\,\big]=\int_{\mathbf{Sym}(\mathbb{R}^m)}\left|\det A-\frac{d_mt}{s_m}\right|\boldsymbol{\Gamma}_{u_m,h_m}\big[\,dA\,\big]. \tag{2.3.40}$$

This is where things get tricky. If we are "lucky" and $u_m\geq 0$, then we can us Lemma C.1.2 to reduce the to express the last integral in terms of the better understood one-point correlation function of the Gaussian ensemble $\mathrm{GOE}_m^v=\mathcal{S}^{0,v}$.

Remark 2.3.9. Before we proceed let first investigate if this is a well founded worry. Set

$$q_m:=q_m(\mathfrak{a})=\frac{h_ms_m}{d_m^2}.$$

Then

$$u_m = h_m \left(1 - \frac{1}{q_m}\right) = h_m \frac{q_m - 1}{q_m}$$

Thus, u_m is negative iff $q_m < 1$.

Using (2.3.16), (2.3.17) and (2.3.17) we deduce

$$q_m(\mathfrak{a}) = \frac{m}{m+2} \cdot \underbrace{\frac{I_{m-1}(\mathfrak{a}) I_{m+3}(\mathfrak{a})}{I_{m+1}(\mathfrak{a})^2}}_{=: R_m(\mathfrak{a})}$$

The Cauchy-Schwarz inequality shows that $R_m(\mathfrak{a}) \geq 1$ for any m and any $\mathfrak{a}$ so

$$q_m(\mathfrak{a}) \geq \frac{m}{m+2}, \quad \forall m, \mathfrak{a}. \tag{2.3.41}$$

This shows that for m large q_m is very close to the critical threshold 1. However as shown in Example B.4.1, the quotient $q_m(\mathfrak{a})$ could be < 1. This happens for example if $\mathfrak{a}(t) = e^{-t^2}$. □

We distinguish three cases.

A. $\boxed{u_m > 0 \text{ so that } q_m > 1.}$ Using (2.3.35), (2.3.36) and (2.3.40) we deduce that

$$\nu_{\mathfrak{a}}\big[\, C \,\big] = \big(\, 2\pi d_m \,\big)^{-m/2} \operatorname{vol}\big[\, B \,\big] \int_C \mathbb{E}_{\mathcal{S}_m^{u_m, h_m}} \left[\left| \det(A - \frac{t d_m}{s_m} \right| \right] \boldsymbol{\Gamma}_{s_m}\big[\, dt \,\big]$$

Making the change in variables $A = \sqrt{h_m} A$ we deduce

$$\mathbb{E}_{\mathcal{S}_m^{u_m, h_m}} \left[\left| \det(A - \frac{t d_m}{s_m} \right| \right] = h_m^{m/2} \mathbb{E}_{\mathcal{S}_m^{2\kappa_m, 1}} \left[\left| \det \left(A - \frac{t d_m}{s_m h_m^{1/2}} \right) \right| \right]$$

and

$$\frac{\nu_{\mathfrak{a}}\big[\, C \,\big]}{\operatorname{vol}\big[\, B \,\big]} = \left(\frac{h_m}{2\pi d_m} \right)^{m/2} \int_C \mathbb{E}_{\mathcal{S}_m^{2\kappa_m, 1}} \left[\left| \det \left(A - \frac{t d_m}{s_m h_m^{1/2}} \right) \right| \right] \boldsymbol{\Gamma}_{s_m}\big[\, dt \,\big]$$

$(t = s_m^{1/2} y)$

$$= \left(\frac{h_m}{2\pi d_m} \right)^{m/2} \int_{s_m^{-1/2} C} \mathbb{E}_{\mathcal{S}_m^{2\kappa_m, 1}} \left[\left| \det \left(A - \frac{y d_m}{(s_m h_m)^{1/2}} \right) \right| \right] \boldsymbol{\Gamma}\big[\, dy \,\big]$$

$$= \left(\frac{h_m}{2\pi d_m} \right)^{m/2} \int_{s_m^{-1/2} C} \mathbb{E}_{\mathcal{S}_m^{2\kappa_m, 1}} \left[\left| \det \left(A - \frac{y}{\sqrt{q_m}} \right) \right| \right] \gamma_1\big[\, dy \,\big].$$

For every $c \in \mathbb{R}$ we denote by $\mathcal{R}_c$ the rescaling map $\mathcal{R}_c : \mathbb{R} \to \mathbb{R}$, $t \mapsto ct$. We set

$$\widehat{\nu}_{\mathfrak{a}} = \frac{1}{\operatorname{vol}\big[\, B \,\big]} \big(\, \mathcal{R}_{s_m^{-1/2}} \,\big) \# \nu_{\mathfrak{a}}$$

We deduce that for any Borel subset $C \subset \mathbb{R}$

$$\widehat{\nu}_{\mathfrak{a}}\big[\, C \,\big] = \nu_{\mathfrak{a}}\big[\, s_m^{1/2} C \,\big] = \left(\frac{h_m}{2\pi d_m} \right)^{m/2} \int_C \underbrace{\mathbb{E}_{\mathcal{S}_m^{2\kappa_m, 1}} \left[\left| \det \left(A - \frac{y}{\sqrt{q_m}} \right) \right| \right] \frac{e^{-y^2/2}}{\sqrt{2\pi}}}_{\boldsymbol{\sigma}_m(y)} dy.$$

To keep the presentation clean I will drop from our notation the dependence of $\widehat{\nu}_{\mathfrak{a}}$ on multiplicative constants. Thus $\widehat{\nu}_{\mathfrak{a}} \propto \mu$ means that $\widehat{\nu}_{\mathfrak{a}} = Z\mu$ where Z is some positive constant. This can be determined from the equality

$$C_m(\mathfrak{a}) = Z\mu\big[\,\mathbb{R}\,\big]$$

where $C_m(\mathfrak{a})$ is determined by (2.3.25) and (2.3.26). We define κ_m by the equality

$$2\kappa_m := \frac{q_m - 1}{q_m} = \frac{s_m}{d_m^2},$$

so that $u_m = 2\kappa_m h_m$.

Observe that the density of $\boldsymbol{\sigma}_m(y)$ is given by

$$\boldsymbol{\sigma}_m(y) = \mathbb{E}_{\mathcal{S}_m^{2\kappa_m,1}}\Big[\,\Big|\det\Big(A - \frac{y}{\sqrt{q_m}}\mathbb{1}_m\Big)\Big|\,\Big]\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \tag{2.3.42}$$

$(\tilde{A} = \sqrt{q_m}A)$

$$= q_m^{-\frac{m}{2}}\mathbb{E}_{\mathcal{S}_m^{2\kappa_m q_m, q_m}}\Big[\,\Big|\det\Big(\tilde{A} - y\mathbb{1}_m\Big)\Big|\,\Big]\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$$

$$\stackrel{(C.1.14)}{=} q_m^{-\frac{m}{2}}2^{\frac{3}{2}}(2q_m)^{\frac{m+1}{2}}\Gamma\left(\frac{m+3}{2}\right)(\theta^+_{m+1,q_m} * \gamma_{2\kappa_m q_m})(y)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$$

$$= 2^{\frac{m+4}{2}}q_m^{\frac{1}{2}}\Gamma\left(\frac{m+3}{2}\right)(\theta^+_{m+1,q_m} * \gamma_{2\kappa_m q_m})(y)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}},$$

where

$$\theta^+_{m+1,v}(x) = \rho_{m+1,v}(x)e^{\frac{x^2}{4v}},$$

and $\rho_{n,v}$ denotes the normalized 1-point correlation function described in Appendix C.1. Hence

$$\widehat{\nu}_{\mathfrak{a}}\big[\,dy\,\big] \propto \big(\,\theta^+_{m+1,q_m} * \gamma_{2\kappa_m q_m}\,\big)(y)\boldsymbol{\Gamma}\big[\,dy\,\big]. \tag{2.3.43}$$

Note that $q_m - 1 = 2\kappa_m q_m$ so

$$(\gamma_{2\kappa_m q_m} * \theta^+_{m+1,q_m})(y) = \frac{1}{\sqrt{4\pi\kappa_m q_m}}\int_{\mathbb{R}} e^{-\frac{(y-x)^2}{4\kappa_m q_m}}\rho_{m+1,q_m}(x)e^{\frac{x^2}{4q_m}}dx$$

$$= \frac{1}{\sqrt{4\pi\kappa_m}}\int_{\mathbb{R}} e^{-\frac{(q_m^{-1/2}y-t)^2}{4\kappa_m}}\rho_{m+1,q_m}\big(\,q_m^{1/2}t\,\big)e^{\frac{t^2}{4}}dt = \frac{q_m^{-1/2}}{\sqrt{4\pi\kappa_m}}\int_{\mathbb{R}} e^{-\frac{(q_m^{-1/2}y-t)^2}{4\kappa_m}}\rho_{m+1,1}(t)e^{\frac{t^2}{4}}dt.$$

Hence

$$(\gamma_{2\kappa_m} * \theta^+_{m+1,q_m})(y)e^{-\frac{y^2}{2}} = \frac{q_m^{-1/2}}{\sqrt{4\pi\kappa_m}}\int_{\mathbb{R}} e^{\frac{t^2}{4} - \frac{(q_m^{-1/2}y-t)^2}{4\kappa_m} - \frac{y^2}{2}}\rho_{m+1,1}(t)dt.$$

$$\frac{t^2}{4} - \frac{(q_m^{-1/2}y - t)^2}{4\kappa_m} - \frac{y^2}{2} = \frac{\kappa_m t^2 - (q_m^{-1/2}y - t)^2 - 2\kappa_m y^2}{4\kappa_m}$$

$(q_m^{-1} + 2\kappa_m = 1)$

$$= \frac{-(y - q_m^{-1}t)^2 - \kappa_m t^2}{4\kappa_m^2},$$

so that

$$(\gamma_{2\kappa_m q_m} * \theta^+_{m+1,q_m})(y) = \frac{1}{\sqrt{4\pi\kappa_m q_m}} \int_{\mathbb{R}} e^{-\frac{(y-q_m^{-1/2}t)^2}{4\kappa_m}} \rho_{m+1,1}(t) e^{-t^2/4} dt$$

$$= \frac{1}{\sqrt{4\pi\kappa_m}} \int_{\mathbb{R}} e^{-\frac{(y-s)^2}{4\kappa_m}} \rho_{m+1,1}\big(q_m^{1/2} s\big) e^{-q_m s^2/4} ds = \frac{1}{\sqrt{4\pi\kappa_m q_m}} \int_{\mathbb{R}} e^{-\frac{(y-s)^2}{4\kappa_m}} \rho_{m+1,1/q_m}\big(s\big) e^{-q_m s^2/4} ds.$$

We deduce

$$\widehat{\nu}_{\mathfrak{a}}\big[\, dy\,\big] \propto \theta^-_{m+1,1/q_m} * \gamma_{2\kappa_m}\big(\, y\,\big)\big[\, dy\,\big], \tag{2.3.44}$$

where

$$\theta^-_{m+1,v}(t) = \rho_{m+1,v}(t) e^{-\frac{t^2}{4v}}. \tag{2.3.45}$$

B. $\boxed{u_m = 0 \text{ so that } q_m = 1.}$ This is obtained from the previous case by letting $q_m \searrow 1$ or, equivalently, $\kappa_m \searrow 0$. We have

$$\widehat{\nu}_{\mathfrak{a}}\big[\, dy\,\big] \propto \theta^-_{m+1,1}\big(\, y\,\big)\big[\, dy\,\big], \tag{2.3.46}$$

C. $\boxed{u_m < 0 \text{ so that } q_m < 1.}$ In this case I will follow the same strategy as in [**111**]. We modify the original function by adding an independent random quantity to it

$$\check{\Phi}_{\mathfrak{a}} = X_{c_m} + \Phi_{\mathfrak{a}}$$

where X_{c_m} is a centered Gaussian random variable with variance c_m and independent of $\Phi_{\mathfrak{a}}$. Fix a box $B \subset \mathbb{R}^m$.

Note that $\nabla\check{\Phi}_{\mathfrak{a}} = \nabla\Phi_{\mathfrak{a}}$ and $\mathrm{Hess}_{\check{\Phi}_{\mathfrak{a}}} = \mathrm{Hess}_{\Phi_{\mathfrak{a}}}$. However, the additive constant X_{c_m} affects the critical values. However the discriminant measure $\check{\mathfrak{D}}_{\mathfrak{a}} = \check{\mathfrak{D}}_{\mathfrak{a},B}$ of $\check{\Phi}_{\mathfrak{a}}$ is related to the discriminant measure $\mathfrak{D}_{\mathfrak{a}}$ via the convolution equation

$$\check{\mathfrak{D}}_{\mathfrak{a}} = \boldsymbol{\Gamma}_{c_m} * \mathfrak{D}_{\mathfrak{a}}.$$

Since the Fourier transform of a Gaussian measure is a nowhere zero function, we deduce that the above equality uniquely determines $\mathfrak{D}_{\mathfrak{a}}$ given $\check{\mathfrak{D}}_{\mathfrak{a}}$. The covariance kernels of $\nabla\check{\Phi}_{\mathfrak{a}}$ and $\mathrm{Hess}_{\check{\Phi}_{\mathfrak{a}}}$ coincide with the respective covariance kernels corresponding to $\Phi_{\mathfrak{a}}$. However

$$\mathbb{E}\big[\,\check{\Phi}_{\mathfrak{a}}(\boldsymbol{x})\,\big] = \check{s}_m := c_m + s_m,\ \ \forall \boldsymbol{x} \in \mathbb{R}^m.$$

This changes the distribution of the conditioned Hessian $\overline{\mathrm{Hess}}_{\mathfrak{a}}(\boldsymbol{x})$ to

$$\underbrace{\left(h_m - \frac{d_m^2}{\check{s}_m}\right)}_{=:\check{u}_m} \delta_{ij}\delta_{k\ell} + h_m\big(\,\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}\,\big)\ \ \forall i \leq j,\ k \leq \ell. \tag{2.3.47}$$

Observe that

$$\check{u}_m = h_m\left(1 - \frac{d_m^2}{h_m\check{s}_m}\right).$$

Set

$$\check{q}_m := \frac{h_m\check{s}_m}{d_m^2} = \frac{h_m s_m}{d_m^2} + \frac{h_m c_m}{d_m^2} = q_m + \frac{h_m c_m}{d_m^2}.$$

Since $q_m < 1$ we can choose c_m such that $\check{q}_m = 1$. More precisely, we let

$$c_m = (1-q_m)d_m^2 h_m^{-1}. \tag{2.3.48}$$

In this case $\check{q}_m = 1$. For any Borel set $C \subset \mathbb{R}$ we set

$$\check{\mu}_{\mathfrak{a}}\big[\, C\,\big] := \mathbb{E}\big[\,\check{\mathfrak{D}}_{\mathfrak{a}}\big[\, C\,\big]\,\big]. \tag{2.3.49}$$

We deduce as in case **B** that

$$\frac{1}{\operatorname{vol}\big[\,B\,\big]}\mathcal{R}_{\check{s}_m^{-1/2}}\check{\mu}_{\mathfrak{a}} \propto \theta^-_{m+1,1}\big(\,y\,\big)\big[\,dy\,\big], \tag{2.3.50}$$

The three cases discussed above can be compactly described by a single statement.

Theorem 2.3.10. *We set*

$$q_m := \frac{h_m s_m}{d_m^2},\ \ r_m := \max\big(\,1, q_m\,\big) \in [1,\infty),$$

$$\check{c}_m := \max\big(\,0, (1-q_m)d_m^2 h_m^{-1}\,\big),\ \ \check{s}_m := s_m + \check{c}_m,\ \ 2\kappa_m := \frac{r_m - 1}{r_m}.$$

We set $\check{\Phi}_{\mathfrak{a}} := X_{\check{c}_m} + \Phi_{\mathfrak{a}}$, *where* $X_{\check{c}_m}$ *is a centered Gaussian variable with variance* $\check{c}_m$, *and independent of* $\Phi_{\mathfrak{a}}$.

For a fixed box $B \subset \mathbb{R}^n$, *denote by* $\check{\mathfrak{D}}_{\mathfrak{a}}$ *the discriminant measure of* $\check{\Phi}_{\mathfrak{a}}\big|_B$. *For any Borel subset* $C \subset \mathbb{R}$ *define*

$$\check{\mu}_{\mathfrak{a}}[C] = \mathbb{E}\big[\,\check{\mathfrak{D}}_{\mathfrak{a}}\big[\,C\,\big]\,\big].$$

Then the following hold.

(i) *The correspondence* $C \mapsto \check{\mu}_{\mathfrak{a}}[C]$ *is a finite Borel measure on* $\mathbb{R}$ *and* $\check{\mu}_{\mathfrak{a}}\big[\,\mathbb{R}\,\big] = C_m(\mathfrak{a})\operatorname{vol}\big]\,B\,\big]$, *where* $C_m(\mathfrak{a})$ *is determined by (2.3.25) and (2.3.26).*

(ii) *If we set*

$$\widehat{\nu}_{\mathfrak{a},m} := \frac{1}{\operatorname{vol}\big[\,B\,\big]}\mathcal{R}_{\check{s}_m^{-1/2}}\check{\mu}_{\mathfrak{a}},$$

then,

$$\widehat{\nu}_{\mathfrak{a}} \propto \theta^-_{m+1,1/r_m} * \gamma_{2\kappa_m}\big(\,y\,\big)\big[\,dy\,\big], \tag{2.3.51}$$

where

$$\theta^-_{m+1,v}(t) = \rho_{m+1,v}(t)e^{-\frac{t^2}{4v}}.$$

□

Remark 2.3.11. (a) Suppose $q_m < 1$ and c_m is given by (2.3.48). Then for any Borel set $C \subset \mathbb{R}$ we have

$$\check{\mu}_{\mathfrak{a}}\big[\,C\,\big] = \int_{\mathbb{R}} \mathbb{E}\big[\,\check{\mathfrak{D}}_{\mathfrak{a}}[C]\,\big|X_{c_m} = x\,\big] = \mu_{\mathfrak{a}} * \boldsymbol{\Gamma}_{c_m}\big[\,C\,\big],$$

so that

$$\check{\mu}_{\mathfrak{a}} = \mu_{\mathfrak{a}} * \boldsymbol{\Gamma}_{c_m}.$$

Passing to Fourier transforms we see that the above equality determines $\mu_{\mathfrak{a}}$ uniquely in terms of $\check{\mu}_{\mathfrak{a}}$. In this case

$$1 = r_m = \frac{\check{s}_m h_m}{d_m^2} > q_m = \frac{s_m h_m}{d_m^2}$$

so

$$1 < \frac{\check{s}_m}{s_m} < \frac{1}{q_m} \overset{(2.3.41)}{\leq} \frac{m+2}{m}.$$

Thus, for m large $\check{s}_m$ and s_m are roughly of same size.

(b) When $q_m > 1$, then $\kappa_m = 1$, $\check{s}_m = s_m$ and $\mathcal{R}_{s_m^{-1/2}}\mu_{\mathfrak{a}} \propto \operatorname{vol}\big[\,B\,\big]\rho_{m+1,1}(t)e^{-\frac{t^2}{4}}$. □

In turns out that, under sufficiently general conditions, the probability measures

$$\bar{\nu}_{\mathfrak{a},m} := \frac{1}{C_m(\mathfrak{a})}\widehat{\nu}_{\mathfrak{a},m}.$$

resemble Gaussian measures for large m.

Theorem 2.3.12. *Suppose that* $r_m = \max\left(1, \frac{s_m h_m}{d_m^2}\right) = \max\left(1, q_m\right)$ *has a limit as* $m \to \infty$,

$$r = \lim_{m\to\infty} r_m \in [1, \infty].$$

Then, as $m \to \infty$, $\bar{\nu}_{\mathfrak{a},m}$ *converges weakly to the Gaussian measure* $\boldsymbol{\Gamma}_{\frac{r+1}{r}}$, *where* $\frac{r+1}{r} = 1$ *if* $r = \infty$.

Proof. Set $\bar{\nu}_m := \bar{\nu}_{\mathfrak{a},m}$. We have

$$\bar{\nu}_m = \frac{1}{K_m}\theta^-_{m+1,\frac{1}{r_m}} * \gamma_{\frac{r_m-1}{r_m}} dy, \tag{2.3.52}$$

where

$$\theta^-_{m+1,\frac{1}{r_m}}(\lambda) = \rho_{m+1,\frac{1}{r_m}}(\lambda)e^{-\frac{r_m\lambda^2}{4}},$$

and

$$K_m = \int_{\mathbb{R}} \theta^-_{m+1,\frac{1}{r_m}} * \gamma_{\frac{r_m-1}{r_m}}(y)dy = \int_{\mathbb{R}} \theta^-_{m+1,\frac{1}{r_m}}(\lambda)d\lambda = \int_{\mathbb{R}} \rho_{m+1,\frac{1}{r_m}}(\lambda)e^{-\frac{r_m\lambda^2}{4}}d\lambda.$$

We set

$$\overline{R}_m(x) := \rho_{m+1,\frac{1}{m}}(x), \quad R_\infty(x) := \frac{1}{2\pi}\boldsymbol{I}_{\{|x|\leq 2\}}\sqrt{4-x^2}.$$

Fix $c \in (0,2)$. In Proposition C.1.4 we proved that

$$\lim_{m\to\infty}\sup_{|x|\leq c}\left|\overline{R}_m(x) - R_\infty(x)\right| = 0, \tag{2.3.53a}$$

and

$$\sup_{|x|\geq c}\left|\overline{R}_m(x) - R_\infty(x)\right| = O(1) \text{ as } m \to \infty. \tag{2.3.53b}$$

Then

$$\rho_{m+1,\frac{1}{r_m}}(\lambda) = \sqrt{\frac{r_m}{m}}\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right), \quad \theta^-_{m+1,\frac{1}{r_m}}(\lambda) = \sqrt{\frac{r_m}{m}}\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right)e^{-\frac{r_m\lambda^2}{4}}.$$

We now distinguish two cases.

Case 1. $r = \lim_{m\to\infty} r_m < \infty$. In particular, $r \in [1,\infty)$. In this case we have

$$K_m = \sqrt{\frac{r_m}{m}}\int_{\mathbb{R}}\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right)e^{-\frac{r_m\lambda^2}{4}}d\lambda,$$

and using (2.3.53a)-(2.3.53b) we deduce

$$\lim_{m\to\infty}\int_{\mathbb{R}}\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right)e^{-\frac{r_m\lambda^2}{4}}d\lambda = R_\infty(0)\int_{\mathbb{R}}e^{-\frac{r\lambda^2}{4}}dr = R_\infty(0)\sqrt{\frac{4\pi}{r}}.$$

Hence

$$K_m \sim K'_m = R_\infty(0)\sqrt{\frac{4\pi}{m}} \text{ as } m \to \infty. \tag{2.3.54}$$

Now observe that

$$\frac{1}{K'_m}\theta^-_{m+1,\frac{1}{r_m}}(\lambda)\big[\,d\lambda\,\big] = \frac{1}{R_\infty(0)}\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right)\frac{r_m}{\sqrt{4\pi}}e^{-\frac{r_m\lambda^2}{4}}d\lambda$$

$$= \frac{1}{R_\infty(0)}\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big].$$

Using (2.3.53a) and (2.3.53b) we conclude that the sequence of measures

$$\frac{1}{K'_m}\theta^-_{m+1,\frac{1}{r_m}}(\lambda)\big[\,d\lambda\,\big]$$

converges weakly to the Gaussian measure $\gamma_{\frac{2}{r}}$. Using this and the asymptotic equality (2.3.54) in (2.3.52) we deduce

$$\lim_{m\to\infty}\bar{\nu}_m = \gamma_{\frac{2}{r}} * \gamma_{\frac{r-1}{r}} = \gamma_{\frac{r+1}{r}}.$$

Case 2. $\lim_{m\to\infty} r_m = \infty$. In this case we have

$$\theta^-_{m+1,\frac{1}{r_m}}(\lambda)\big[\,d\lambda\,\big] = \sqrt{\frac{4\pi}{m}}\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big].$$

Lemma 2.3.13. *The sequence of measures*

$$\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right)\gamma_{\frac{2}{r_m}}d\lambda$$

converges weakly to the measure $R_\infty(0)\delta_0$.

Proof. Fix a bounded continuous function $f:\mathbb{R}\to\mathbb{R}$. Observe first that

$$\lim_{m\to\infty}\underbrace{\int_{\mathbb{R}}\left(\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right) - R_\infty\left(\sqrt{\frac{r_m}{m}}\lambda\right)\right)f(\lambda)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big]}_{=D_m} = 0. \tag{2.3.55}$$

Indeed, we have

$$D_m = \underbrace{\int_{|\lambda|<c\frac{\sqrt{m}}{\sqrt{r_m}}}\left(\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right) - R_\infty\left(\sqrt{\frac{r_m}{m}}\lambda\right)\right)f(\lambda)\gamma_{\frac{2}{r_m}}\big[\,d\lambda\,\big]}_{=:D'_m}$$

$$+\underbrace{\int_{|\lambda|>c\frac{\sqrt{m}}{\sqrt{r_m}}}\left(\overline{R}_m\left(\sqrt{\frac{r_m}{m}}\lambda\right) - R_\infty\left(\sqrt{\frac{r_m}{m}}\lambda\right)\right)f(\lambda)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big]}_{=:D''_m}.$$

Observe that

$$D'_m \le \sup_{|x|\le c}\big|\overline{R}_m(x) - R_\infty(x)\big|\int_{|\lambda|<c\frac{\sqrt{m}}{\sqrt{r_m}}}f(\lambda)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big]$$

and invoking (2.3.53a) we deduce

$$\lim_{m\to\infty}D'_m = 0.$$

Using (2.3.53b) we deduce that there exists a constant $S>0$ such that

$$D'_m \le S\int_{|\lambda|>c\frac{\sqrt{m}}{\sqrt{r_m}}}\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big].$$

On the other hand, Chebyshev's inequality shows that

$$\int_{|\lambda|>c\frac{\sqrt{m}}{\sqrt{r_m}}}\gamma_{\frac{2}{r_m}}\big[\,d\lambda\,\big]\leq\frac{2}{c^2m}.$$

Hence

$$\lim_{m\to\infty}D''_m=0.$$

This proves (2.3.55).

The sequence of measures $\gamma_{\frac{2}{r_m}}(\lambda)d\lambda$ converges to δ_0 so that

$$R_\infty(0)f(0)=\lim_{m\to\infty}\int_{\mathbb{R}}R_\infty(0)f(\lambda)\gamma_{\frac{2}{r_m}}(\lambda)d\lambda.$$

Using (2.3.55) and the above equality we deduce that the conclusion of the lemma is equivalent to

$$\lim_{m\to\infty}\underbrace{\int_{\mathbb{R}}\left(R_\infty(0)-R_\infty\left(\sqrt{\frac{r_m}{m}}\lambda\right)\right)f(\lambda)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big]}_{=F_m}=0.\tag{2.3.56}$$

To prove this we decompose F_m as follows.

$$F_m=\underbrace{\int_{|\lambda|<m^{-\frac{1}{4}}\frac{\sqrt{m}}{\sqrt{r_m}}}\left(R_\infty(0)-R_\infty\left(\sqrt{\frac{r_m}{m}}\lambda\right)\right)f(\lambda)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big]}_{=:F'_m}$$

$$+\underbrace{\int_{|\lambda|>m^{-\frac{1}{4}}\frac{\sqrt{m}}{\sqrt{r_m}}}\left(R_\infty(0)-R_\infty\left(\sqrt{\frac{r_m}{m}}\lambda\right)\right)f(\lambda)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big]}_{=:F''_m}.$$

Observe that

$$F'_m\leq\sup_{|x|\leq m^{-\frac{1}{4}}}\big|\,R_\infty(0)-R_\infty(x)\,\big|\int_{|\lambda|<m^{-\frac{1}{4}}\frac{\sqrt{m}}{\sqrt{r_m}}}f(\lambda)\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big].$$

Since R_∞ is continuous at 0 we deduce

$$\lim_{m\to\infty}F'_m=0.$$

Since R_∞ and f are bounded we deduce that there exists a constant $S>0$ such that

$$F''_m\leq S\int_{|\lambda|>m^{-\frac{1}{4}}\frac{\sqrt{m}}{\sqrt{r_m}}}\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big].$$

On the other hand, Chebyshev's inequality shows that

$$\int_{|\lambda|>m^{-\frac{1}{4}}\frac{\sqrt{m}}{\sqrt{r_m}}}\boldsymbol{\Gamma}_{\frac{2}{r_m}}\big[\,d\lambda\,\big]\leq\frac{2}{\sqrt{m}}.$$

Hence

$$\lim_{m\to\infty}F''_m=0.$$

This proves (2.3.56) and the lemma. □

Lemma 2.3.13 shows that

$$K_m \sim K'_m = \sqrt{\frac{4\pi}{m}} R_\infty(0),$$

and

$$\lim_{m\to\infty} \frac{1}{K_m} \theta^-_{m+1,\frac{1}{r_m}}(\lambda)\big[\,d\lambda\,\big] = \delta_0\big[\,\lambda\,\big].$$

On the other hand,

$$\lim_{m\to\infty} \gamma_{\frac{r_{m-1}}{r_m}}\big[\,d\lambda\,\big] = \boldsymbol{\Gamma}\big[\,d\lambda\,\big],$$

so that

$$\lim_{m\to\infty} \bar{\nu}_m = \delta_0 * \gamma_1 = \gamma_1.$$

□

Remark 2.3.14. Note that for any Borel subset $C \subset \mathbb{R}$, the number $\bar{\nu}_{\mathfrak{a},m}\big[\,C\,\big]$ is the expected proportion of critical values of

$$\frac{1}{\sqrt{\check{s}_m}}\big(\,\Phi_{\mathfrak{a}} + X_{\check{c}_m}\,\big)\big|_B$$

located in C. For large m the bulk of these critical values are located in an interval of size $O(1)$ centered at the origin. Thus, the bulk of critical values of $\Phi_{\mathfrak{a}} + X_{\check{c}_m}$ is located in an interval of size $O\big(\sqrt{s_m}\,\big)$ centered at the origin. Recall that

$$s_m = s_m(\mathfrak{a}) = \frac{1}{(2\pi)^{m/2}\Gamma(m/2)} I_{m-1}(\mathfrak{a})$$

The large m behavior is sensitive to the choice of amplitude; see Appendix B.4. For example, if $\mathfrak{a}(t) = e^{-t^2/2}$, then

$$s_m(\mathfrak{a}) = \frac{1}{(2\pi)^{m/2}}.$$

However if

$$\mathfrak{a}(t) = \exp\Big(\,-\frac{1}{2}\log(t)\log(\log t)\,\Big), \;\; \forall t > 1,$$

then

$$\log s_m(\mathfrak{a}) \sim e^{m-1} \;\text{ as } m \to \infty.$$

□

2.3.4. A probabilistic computation of a Mehta integral.

Recall that GOE_n^v, $v > 0$ is the Gaussian ensemble of symmetric $n \times n$ matrices $A = (a_{ij})_{1\le i,j\le n}$ where the entries $\big(\,a_{ij}\,\big)_{1\le i\le j\le n}$ are independent centered Gaussian variables with variances

$$\mathbb{E}\big[\,a_{ii}^2\,\big] = 2v, \;\; \mathbb{E}\big[\,a_{jk}^2\,\big] = v, \;\; \forall i, \;\; \forall j < k. \tag{2.3.57}$$

As detailed in Appendix C.1, the *normalized* 1-point correlation function $\rho_{n,v}(x)$ of GOE_n^v is the function $\rho_{n,v} : \mathbb{R} \to [0,\infty)$ uniquely determined by the equality

$$\frac{1}{n}\mathbb{E}_{\mathrm{GOE}_n^v}\big[\,\mathrm{tr}\, f(X)\,\big] = \int_{\mathbb{R}} f(\lambda)\rho_{n,v}(\lambda)d\lambda,$$

for any bounded Borel measurable $f : \mathbb{R} :\to \mathbb{R}$. For example, if $f = \boldsymbol{I}_B$, $B \subset \mathbb{R}$, then

$$\int_B \rho_{n,v}(\lambda)d\lambda$$

is the expected fraction of eigenvalues of a random symmetric matrix X located in B.

One can show that for any bounded Borel measurable function function $F : \mathbf{Sym}(\mathbb{R}^n) \to \mathbb{R}$ that is invariant with respect to the conjugation by orthogonal transformations we have

$$\mathbb{E}_{\mathrm{GOE}_n^v}\big[\,F(X)\,\big] = \frac{1}{\boldsymbol{Z}_n(v)}\int_{\mathbb{R}^m} F\big(\,\mathrm{Diag}(\lambda_1,\dots,\lambda_n)\,\big)\underbrace{\left(\prod_{1\le i<j\le n}|\lambda_i-\lambda_j|\right)\prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}}_{=:Q_{n,v}(\lambda)}|d\lambda_1\cdots d\lambda_n|.$$

Then

$$\rho_{n,v}(x) = \frac{1}{\boldsymbol{Z}_n(v)}\int_{\mathbb{R}^{n-1}} Q_{n,v}(x,\lambda_2,\dots,\lambda_n)d\lambda_1\cdots d\lambda_n.$$

$$\boldsymbol{Z}_m = \boldsymbol{Z}_m(1/2) := \int_{\mathbb{R}^m}\prod_{1\le i<j\le m}|\lambda_i-\lambda_j|\prod_{i=1}^m e^{-\frac{\lambda_i^2}{2}}|d\lambda_1\cdots d\lambda_m|.$$

The integral in the right-hand-side is known as *Mehta integral*. I will prove that

$$\boldsymbol{Z}_m = (2\pi)^{\frac{m}{2}}\prod_{j=0}^{m-1}\frac{\Gamma(\frac{j+3}{2})}{\Gamma(3/2)} = 2^{\frac{3m}{2}}\prod_{j=0}^{m-1}\Gamma\Big(\frac{j+3}{2}\Big). \tag{2.3.58}$$

This equality was first proved in 1960 by M. L. Mehta, [**98**]. Later Mehta observed that this integral was known earlier to N. G. de Brujin [**28**]. It was subsequently observed that Mehta's integral is a limit of the *Selberg integrals*, [**5**, Eq. (2.5.11)], [**61**, Sec. 4.7.1].

The goal of this subsection is to provide a probabilistic computation of the Mehta integra.I follow the approach in [**120**]. The strategy is easy to describe. We argue inductively. An immediate direct computation shows that

$$\boldsymbol{Z}_1 = \int_{\mathbb{R}} e^{-t^2/2}dt = (2\pi)^{1/2}.$$

To compute the ratio $\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}$ we observe that the eigenvalues of $A\in\mathbf{Sym}(\mathbb{R}^{m+1})$ coincide with the critical values of the restriction to the unit sphere of the quadratic function $\boldsymbol{x}\mapsto(A\boldsymbol{x},\boldsymbol{x})$. The Kac-Rice formula will provide a description of the mean distribution of these critical values which will lead to an explicit evaluation of $\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}$. Here are the details.

For each $A\in\mathbf{Sym}(\mathbb{R}^{m+1})$ we obtain a quadratic function

$$q_A:\mathbb{R}^{m+1}\to\mathbb{R},\;\; q_A\big(\,\boldsymbol{x}\,\big) = \frac{1}{2}\big(\,A\boldsymbol{x},\boldsymbol{x}\,\big).$$

We denote by Φ_A its restriction to the unit sphere

$$S^m = \big\{\,\boldsymbol{x}\in\mathbb{R}^{m+1};\;\;\|\boldsymbol{x}\|=1\,\big\}$$

Above, $(-,-)$ and $\|-\|$ denote respectively the canonical inner product and its associated norm on $\mathbb{R}^{m+1}$.

When A runs in the Gaussian ensemble GOE_{m+1}^v we obtain a Gaussian function

$$\Phi=\Phi_A:S^m\to\mathbb{R}.$$

This is invariant under the natural $O(m+1)$-action on S^m. As shown in [**108**, Ex.1.20], the function Φ_A is Morse for generic A, where genericity is understood in Baire categoric terms.

Lemma 2.3.15. *The Gaussian function Φ_A is* a.s. *Morse.*

Proof. It suffices to show that the Gaussian section $\nabla\Phi_A$ of TS^m is ample. Let $\boldsymbol{x} \in S^m$. If

$$\mathrm{Proj}_{\boldsymbol{x}} : \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$$

the orthogonal projection onto $T_{\boldsymbol{x}}S^m$, then

$$\nabla\Phi_A(\boldsymbol{x}) = \mathrm{Proj}_{\boldsymbol{x}} A\boldsymbol{x} = A\boldsymbol{x} - (\boldsymbol{x}, \boldsymbol{x})\boldsymbol{x}.$$

The map

$$\mathbf{Sym}(\mathbb{R}^{m+1}) \ni A \mapsto A\boldsymbol{x} \in \mathbb{R}^{m+1}$$

is onto and thus the map

$$\mathbf{Sym}\left(\mathbb{R}^{m+1}\right) \ni A \mapsto \mathrm{Proj}_{\boldsymbol{x}} A\boldsymbol{x} \in T_{\boldsymbol{x}}S^m$$

is also onto, thus proving that the gradient $\nabla\Phi_A(\boldsymbol{x})$ is nondegenerate since the Gaussian ensemble $\mathrm{GOE}_{m+1,v}$ is nondegenerate. □

Consider the *spectral measure* of A,

$$\boldsymbol{\sigma}_A := \sum_{\lambda \in \mathrm{Spec}(A)} \mathrm{mult}(\lambda)\delta_\lambda.$$

The discriminant measure of $2\Phi_A$ is

$$\boldsymbol{D}_A = \sum_{\nabla\Phi_A(\boldsymbol{x})=0} \delta_{2\Phi_A(x)} = (2\Phi_{\mathfrak{a}})_{\#}\mathfrak{C}[-, \Phi_A]$$

The critical values of $2\Phi_A$ are precisely the eigenvalues of A and the critical points are the unit eigenvectors of A. The function is Morse iff A is simple, i.e., its eigenvalues are distinct. In this case to each critical value of A there corresponds exactly *two* critical points. Hence

$$\boldsymbol{D}_A = 2\boldsymbol{\sigma}_A \ \text{ a.s..}$$

Then for any Borel subset $C \subset \mathbb{R}$ we have

$$\frac{1}{(m+1)}\mathbb{E}\big[\,\mathfrak{D}_A\big[\,C\,\big]\,\big] = \frac{2}{m+1}\mathbb{E}\big[\,\boldsymbol{\sigma}_A\big[\,C\,\big]\,\big] = 2\int_C \rho_{m+1,v}(\lambda)d\lambda. \tag{2.3.59}$$

We will determine $\mathbb{E}\big[\,\mathfrak{D}_A\big[\,C\,\big]\,\big]$ using the Kac-Rice formula (2.2.17) .

For $\boldsymbol{x} \in S^m$ and $A \in \mathrm{GOE}^v_{m+1}$ we denote by $\mathrm{Hess}_A(\boldsymbol{x})$ the Hessian of Φ_A at $\boldsymbol{x}$ viewed as a symmetric operator $T_{\boldsymbol{x}}S^m \to T_{\boldsymbol{x}}S^m$. Here the Hessian is defined in the sense of Riemann geometry

$$\mathrm{Hess}_A = \nabla(\nabla\Phi_A),$$

where $\nabla\Phi_A$ is the gradient of Φ_A, i.e., the metric dual of the differential $d\Phi_A$, and ∇ denotes the Levi-Civita connection of the round metric.

Denote by $(x^0, x^1, \dots, x^m)$ the canonical Euclidean coordinates on $\mathbb{R}^{m+1}$ so

$$A = \big(\,a_{ij}\,\big)_{0\le i,j\le m},\ \ \mathbb{E}\big[\,a_{ii}\,\big] = 2v,\ \ \mathbb{E}\big[\,a_{ij}\,\big| = v,\ \ i \neq j,\ \ \mathbb{E}\big[\,a_{ij}a_{k\ell}\,\big] = 0,\ \ (i,j) \neq (k,\ell).$$

Since Φ_A is $O(m+1)$ invariant, the distribution of the random operator $\mathrm{Hess}_A(\boldsymbol{x})$ is independent of $\boldsymbol{x}$ so it suffices to determine it at any point of our choosing. Suppose that $\boldsymbol{x}$ is the north pole

$$\boldsymbol{x} = \boldsymbol{n} = (1, 0, \dots, m) \in \mathbb{R}^{m+1}.$$

Then $T_{\boldsymbol{n}}S^m = \{x^0 = 0\}$ and $\boldsymbol{x}_* := \big(\,x^1, \dots, x^m\big)$ are orthonormal coordinates on $T_{\boldsymbol{n}}S^m$. The next result describes the statistics of $\nabla\Phi_A(\boldsymbol{n})$ and $\mathrm{Hess}_A(\boldsymbol{n})$.

Lemma 2.3.16. *The Gaussian vector $\nabla\Phi_A(\boldsymbol{n})$ is given in the orthonormal coordinates $\boldsymbol{x}^*$ on $T_{\boldsymbol{n}}S^m$ by*

$$\Phi_A(\boldsymbol{n}) = \frac{1}{2}a_{00}, \;\; \nabla\Phi_A(\boldsymbol{n}) = \begin{bmatrix} \partial_{x^1}\Phi_A(\boldsymbol{n}) \\ \vdots \\ \partial_{x^m}\Phi_A(\boldsymbol{n}) \end{bmatrix} = \begin{bmatrix} a_{01} \\ \vdots \\ a_{0m} \end{bmatrix} \tag{2.3.60}$$

and thus it has mean zero and variance operator $v\mathbb{1}_m$. Moreover, in the orthonormal coordinates $\boldsymbol{x}^*$ on $T_{\boldsymbol{n}}S^m$ we have

$$\operatorname{Hess}_A(\boldsymbol{n}) = A_* - a_{00}\mathbb{1}_m, \;\; A_* = (a_{ij})_{1\leq i\leq m}.$$

In particular, $\operatorname{Hess}_A(\boldsymbol{n}) \in \mathcal{S}_m^{2v,v}$, where $\mathcal{S}_m^{u,v}$ is the $O(m)$-invariant Gaussian ensemble defined by (C.1.3).

Proof. The coordinates $\boldsymbol{x}_*$ also define local coordinates on S^m. More precisely, the upper hemisphere

$$S^m_+ := \left\{ \boldsymbol{x} \in S^m; \;\; x^0 > 0 \right\}$$

admits the parametrization

$$\boldsymbol{x}^* \mapsto \boldsymbol{x}\left(\boldsymbol{x}_*\right) = \left(x^0(\boldsymbol{x}_*), \boldsymbol{x}^*\right) \in S^m, \;\; x^0(\boldsymbol{x}_*) = \sqrt{1 - \|\boldsymbol{x}_*\|^2}.$$

We deduce that in the coordinates $\boldsymbol{x}^*$ the round metric on S^m satisfies

$$g_{ij} = \delta_{ij} + O\left(\|\boldsymbol{x}_*\|^2\right) \text{ near } \boldsymbol{n}. \tag{2.3.61}$$

This proves that in the coordinates $\boldsymbol{x}^*$ the Christoffel symbols vanish at $\boldsymbol{n}$. This has two consequences: the gradient $\nabla\Phi_A(\boldsymbol{n})$ is given by the vector $\left(\partial_{x^i}\Phi_A(\boldsymbol{n})\right)_{1\leq i\leq m}$ and the Hessian $\operatorname{Hess}_A(\boldsymbol{n})$ is given by the matrix $\left(\partial^2_{x^ix^j}\Phi_A(\boldsymbol{n})\right)_{1\leq i,j\leq m}$.

On the upper hemisphere we will view Φ_A as a function of $\boldsymbol{x}_*$. Since

$$x^0 = 1 - \frac{1}{2}\|\boldsymbol{x}^*\|^2 + O\left(\|\boldsymbol{x}^*\|^4\right) \tag{2.3.62}$$

we deduce that in the coordinates $\boldsymbol{x}_*$ we have

$$\begin{aligned} \Phi_A(\boldsymbol{x}) &= \frac{1}{2}a_{00}\left(1 - \|\boldsymbol{x}_*\|^2\right) + \frac{1}{2}\sum_{j=1}^m a_{jj}(x^j)^2 + \sum_{0\leq j<k\leq m} a_{jk}x^jx^k + O\left(\|\boldsymbol{x}^*\|^4\right) \\ &= \frac{1}{2}a_{00} + \frac{1}{2}\sum_{j=1}^m \left(a_{jj} - a_{00}\right)(x^j)^2 + \sum_{0\leq j<k\leq m} a_{jk}x^jx^k + O\left(\|\boldsymbol{x}^*\|^4\right) \quad (2.3.63) \\ &\overset{(2.3.62)}{=} \frac{1}{2}a_{00} + \sum_{k=1}^m a_{0k}x^k + \frac{1}{2}\sum_{j=1}^m \left(a_{jj} - a_{00}\right)(x^j)^2 + \sum_{1\leq j<k\leq m} a_{jk}x^jx^k + O\left(\|\boldsymbol{x}^*\|^3\right). \end{aligned}$$

We deduce that $\Phi_A(\boldsymbol{n}) = \frac{1}{2}a_{00}$ and $\partial_{x^i}\Phi_A(\boldsymbol{n}) = a_{0i}$, $1 \leq i \leq m$. This proves (2.3.60).

Note that $A_* \in \mathrm{GOE}^v_m$. Using (2.3.63) we deduce that

$$\operatorname{Hess}_A(\boldsymbol{n}) = A_* - a_{00}\mathbb{1}_m.$$

Since a_{00} is independent of A_* we deduce from (C.1.5) that $\operatorname{Hess}_A(\boldsymbol{n}) \in \mathcal{S}_m^{2v,v}$, where $\mathcal{S}_m^{u,v}$ is the $O(m)$-invariant Gaussian ensemble defined by (C.1.3). □

The covariance kernel of Φ_A is

$$\mathcal{K}_A(\boldsymbol{n},\boldsymbol{x}) = \mathbb{E}\big[\,\Phi_A(\boldsymbol{n})\Phi_A(\boldsymbol{x})\,\big] = \frac{1}{4}\big(\,1-\|\boldsymbol{x}_*\|^2\,\big)\mathbb{E}\big[\,a_{00}^2\,\big] = \frac{v}{2}\big(\,1-\|\boldsymbol{x}_*\|^2\,\big).$$

Since $\Phi_A(\boldsymbol{n}) = \frac{1}{2}a_{00}$ we deduce from (2.3.60) that $\nabla\Phi_A(\boldsymbol{n}) = (a_{01},\dots,a_{0n})$ is independent of $\Phi_A(\boldsymbol{n})$. Lemma 2.3.16 and (2.3.60) imply that $\nabla\Phi_A(\boldsymbol{n})$ and $\mathrm{Hess}_A(\boldsymbol{n})$ are also independent.

If we set

$$L_{ij} := \ell_{ij}\big(\,\mathrm{Hess}_A(\boldsymbol{n})\,\big),\;\; \Omega_{ij} := \omega_{ij}\big(\,\mathrm{Hess}_A(\boldsymbol{n})\,\big) = \begin{cases} L_{ii}, & i=j,\\ \sqrt{2}L_{ij}, & i<j\end{cases}, \tag{2.3.64}$$

where ℓ_{ij} and ω_{ij} are defined by (2.3.21), then

$$\mathbb{E}\big[\,L_{ij}L_{k\ell}\,\big] = 2v\delta_{ij}\delta_{k\ell} + v(\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk}),\;\; \forall 1\le i,j,.k,\ell\le m. \tag{2.3.65}$$

Set

$$W := \begin{bmatrix} \Phi_A(\boldsymbol{n})\\ \nabla\Phi_A(\boldsymbol{n})\end{bmatrix} = \begin{bmatrix} \frac{1}{2}a_{00}\\ a_{01}\\ \vdots\\ a_{0m}\end{bmatrix}.$$

Note that

$$\mathrm{Var}\big[\,W\,\big] = \mathrm{Diag}\,\big(\,\frac{v}{2},\underbrace{2v,\dots,2v}_{m}\,\big). \tag{2.3.66}$$

Clearly the matrix $\mathrm{Var}\big[\,W\,\big]$ is invertible, proving Φ_A is J_1-ample.

Denote by $\overline{\mathrm{Hess}}_A(\boldsymbol{n})$ the random symmetric matrix with variance given by the regression formula (1.1.18)

$$\begin{aligned}\mathrm{Var}\big[\overline{\mathrm{Hess}}_A(\boldsymbol{n})\big] &= \mathrm{Var}\big[\,\mathrm{Hess}_A(\boldsymbol{n})\,\big]\\ &- \mathrm{Cov}\big[\,\mathrm{Hess}_A(\boldsymbol{n}),W\,\big]\,\mathrm{Var}\big[\,W\,\big]^{-1}\,\mathrm{Cov}\big[\,W,\mathrm{Hess}_A(\boldsymbol{n})\,\big].\end{aligned}$$

Set

$$\bar{L}_{ij} := \ell_{ij}\big(\overline{\mathrm{Hess}}_A(\boldsymbol{n})\,\big),\;\; \bar{\Omega}_{ij} := \omega_{ij}\big(\overline{\mathrm{Hess}}_A(\boldsymbol{n})\,\big),$$

and

$$C_{ij|k} := \mathrm{Cov}\big[\,\Omega_{ij},W_k\,\big],\;\; 1\le i\le j\le m,\;\; 0\le k\le m.$$

Note that

$$C_{ij|k} = 0,\;\; \forall i,j,\;\; \forall k>0,\;\; C_{ij|k}=0,\;\; \forall i<j,\;\; \forall k\ge 0,$$

and

$$C_{ii|0} = \frac{1}{2}\mathbb{E}\big[\,(a_{ii}-a_{00})a_{00}\,\big] = -\frac{1}{2}\mathbb{E}\big[\,a_{00}^2\,\big] = -v.$$

If we write

$$\mathrm{Var}\big[\,W\,\big]^{-1} = \big(\,t_{ab}\,\big)_{0\le a,b\le m},$$

then

$$\begin{aligned}\mathbb{E}\big[\bar{\Omega}_{ij}\bar{\Omega}_{k\ell}\,\big] &= \mathbb{E}\big[\,\Omega_{ij}(\boldsymbol{x})\Omega_{k\ell}(\boldsymbol{x})\,\big] - \sum_{a,b=0}^{m} C_{ij|a}t_{ab}C_{k\ell|b}\\ &\stackrel{(2.3.66)}{=} \mathbb{E}\big[\,\Omega_{ij}(\boldsymbol{x})\Omega_{k\ell}(\boldsymbol{x})\,\big] - \frac{2}{v}C_{ij|0}C_{k\ell|0}.\end{aligned}$$

We deduce

$$\mathbb{E}\big[\bar{\Omega}_{ii}\bar{\Omega}_{jj}\,\big] = \mathbb{E}\big[\,\Omega_{ii}(\boldsymbol{x})\Omega_{jj}(\boldsymbol{x})\,\big] - 2v = 0,\;\; i\ne j,$$

$$\mathbb{E}\big[\,\overline{\Omega}_{ii}^2\,\big] = 2v,$$
$$\mathbb{E}\big[\,\overline{\Omega}_{ij}\overline{\Omega}_{k\ell}\,\big] = \mathbb{E}\big[\,\Omega_{ij}(\boldsymbol{x})\Omega_{k\ell}(\boldsymbol{x})\,\big],\ \ \forall 1 \leq i < j,\ \ 1 \leq k < \ell.$$

Using (2.3.64) and (2.3.65) we conclude that

$$\overline{\mathrm{Hess}}_A \in \mathrm{GOE}_m^v\,. \tag{2.3.67}$$

The regression operator

$$R_{\mathrm{Hess}_A,W} = \mathrm{Cov}\,\big[\ \mathrm{Hess}_A, W\,\big]\ \mathrm{Var}\,\big[\,W^{-1}\,\big] : \mathbb{R}^{m+1} \to \mathbf{Sym}_m(\mathbb{R})$$

is

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_m \end{bmatrix} \mapsto -2w_0\mathbb{1}_m.$$

We set $H := \overline{\mathrm{Hess}}_A$, so $H \in \mathrm{GOE}_m^v$. Using the regression formula (1.1.19) we deduce

$$\mathbb{E}\big[\,|\,\mathrm{Hess}_A(\boldsymbol{n})|\,\big|\,W(\boldsymbol{n}) = (t/2,0)\,\big] = \mathbb{E}_{\mathrm{GOE}_m^v}\big[\,|\det(H - vt)|\,\big].$$

Since $2\Phi_A(\boldsymbol{n}) = a_{00}$ is Gaussian with variance $2v$, we deduce from the Kac-Rice formula (2.2.17) that for any Borel subset $C \subset \mathbb{R}$ we have

$$\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big] = \int_C \rho_A(t)\boldsymbol{\gamma}_{2v}\big[\,dt\,\big],$$

where

$$\begin{aligned}
\rho_A(t) &= \int_{S^m} \mathbb{E}\big[\,|\,\mathrm{Hess}_A(\boldsymbol{x})|\,\big|\,2\Phi_A(\boldsymbol{x}) = t,\ \nabla\Phi_A(\boldsymbol{x}) = 0\,\big]p_{\nabla\Phi_A(\boldsymbol{x})}(0)d\boldsymbol{x} \\
&\overset{(2.3.60)}{=} \big(\,2\pi v\,\big)^{-m/2}\int_{S^m} \underbrace{\mathbb{E}\big[\,|\,\mathrm{Hess}_A(\boldsymbol{x})|\,\big|\,W(\boldsymbol{x}) = (t/2,0)\,\big]}_{\text{independent of } \boldsymbol{x}}d\boldsymbol{x} \\
&= \big(\,2\pi v\,\big)^{-m/2}\mathbb{E}\big[\,|\,\mathrm{Hess}_A(\boldsymbol{n})|\,\big|\,W(\boldsymbol{n}) = (t/2,0)\,\big]\,\mathrm{vol}\,\big[\,S^m\,\big] \\
&= \big(\,2\pi v\,\big)^{-m/2}\,\mathrm{vol}\,\big[\,S^m\,\big]\mathbb{E}_{\mathrm{GOE}_m^v}\big[\,|\det(H - vt)|\,\big].
\end{aligned} \tag{2.3.68}$$

Lemma C.1.1 shows that

$$\mathbb{E}_{\mathrm{GOE}_{m,v}}\big[\ \det(H - vt)|\,\big] = (2v)^{\frac{m+1}{2}}e^{\frac{v^2t^2}{4v}}\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\rho_{m+1,v}(vt).$$

Assume now that $v = 1$.

$$\mathbb{E}_{\mathrm{GOE}_m^1}\big[\ \det(H - t)|\,\big] = e^{\frac{t^2}{4}}2^{\frac{m+1}{2}}\pi^{-1/2}\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\rho_{m+1,1}(t).$$

Since $\boldsymbol{\gamma}_2\big[\,dt\,\big] = e^{-\frac{t^2}{4}}\frac{dt}{\sqrt{4\pi}}$ we deduce

$$\begin{aligned}
\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big] &= \int_C \rho_A(t)\boldsymbol{\gamma}_{2v}\big[\,dt\,\big] \\
&= \big(\,2\pi\,\big)^{-m/2}2^{\frac{m+1}{2}}\,\mathrm{vol}\,\big[\,S^m\,\big]\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\int_C \rho_{m+1,1}(t)\frac{dt}{\sqrt{4\pi}}.
\end{aligned}$$

On the other hand, we deduce from (2.3.59) that

$$\frac{1}{(m+1)}\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big] = 2\int_C \rho_{m+1,1}(t)dt,$$

so that

$$\frac{\left(2\pi\right)^{-m/2}2^{\frac{m+1}{2}}\operatorname{vol}\left[\,S^m\,\right]}{(m+1)}\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}(4\pi)^{-1/2}=2. \tag{2.3.69}$$

Using the fact that

$$\frac{\operatorname{vol}\left[\,S^m\,\right]}{m+1}=\frac{\pi^{\frac{m+1}{2}}}{\Gamma(\frac{m+3}{2})}$$

we deduce

$$\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}=\frac{\Gamma(\frac{m+3}{2})}{\pi^{\frac{m+1}{2}}}\cdot\frac{2(2\pi)^{m/2}(4\pi)^{1/2}}{2^{\frac{m+1}{2}}}=2^{3/3}\Gamma\Big(\frac{m+3}{2}\Big).$$

Note that

$$\boldsymbol{Z}_1=\int_{\mathbb{R}}e^{-t^2/2}dt=(2\pi)^{1/2}.$$

We deduce immediately the equality (C.1.8)

$$\boldsymbol{Z}_m=\boldsymbol{Z}_1\prod_{j=1}^{m-1}\frac{\boldsymbol{Z}_{j+1}}{\boldsymbol{Z}_j}=2^{\frac{3m}{2}}\prod_{j=0}^{m-1}\Gamma\Big(\frac{j+3}{2}\Big).$$

2.3.5. Random matrices and Morse functions on Grassmannians. The example discussed in the previous subsection is a special manifestation of a more general phenomenon I will discuss in some detail.

2.4. Higher momentums

As in previous section suppose that $F:\mathscr{V}\to\boldsymbol{U}$ is an ample Gaussian C^1-field, where $\boldsymbol{U}$ is a real Euclidean space, $\mathscr{V}$ is an open subset of the Euclidean space $\boldsymbol{V}$ and $\dim\boldsymbol{U}=\dim\boldsymbol{V}=m$.

2.4.1. A preview. For any Borel subset $B\subset\mathscr{V}$ we denote by $Z[B]=Z[B,F]$ the number of zeros of F inside B. The question we will address in this section concerns the finiteness of the various momentums of F. The first nontrivial case has to do with the variance of $Z[B]$. Which conditions on F will guarantee the finiteness of the variance of $Z[B]$ when B is say a box in $\mathscr{V}$?

We approach this using a simple trick. Consider the random field

$$\widehat{F}:\mathscr{V}\times\mathscr{V}\to\boldsymbol{U}\times\boldsymbol{U},\ \ (v_1,v_2)\mapsto\big(\,F(v_1),F(v_2)\,\big).$$

Note that $Z[B]^2=Z\big(\,\widehat{F},B\times B\,\big)$ so we may try to apply the local Kac-Rice formula to the Gaussian field $\widehat{F}$. There is a an immediate obstacle on our way, namely, the Gaussian field $\widehat{F}$ fails to be ample along the diagonal

$$\Delta=\big\{\,(v_1,v_2)\in\mathscr{V}^2;\ \ v_1=v_2\,\big\}$$

since the $\boldsymbol{U}\times\boldsymbol{U}$ Gaussian vector $\big(\,F(v),F(v)\,\big)|$ is *degenerate*!

We are forced to remove the diagonal. We set $\mathscr{V}^2_*:=\mathscr{V}^2\setminus\Delta$, $B^2_*:=B^2\setminus\Delta$. Then

$$Z\big[\,B^2_*,\widehat{F},\,\big]=Z[B]^2-Z[B]=Z[B]\big(\,Z[B]-1\,\big)=\big(\,Z[B]\,\big)_2.$$

Above, for any $x\in\mathbb{R}$ and $k\in\mathbb{N}$ we denote by $(x)_k$ the *falling factorial*[5]

$$(x)_k := x(x-1)\cdots(x-k+1)=\prod_{j=0}^{k-1}(x-j).$$

We will attempt to use the Kac-Rice formula for the random field $\widehat{F}|_{\mathscr{V}_*^2}$. It is not clear yet that this Gaussian field is ample, we hope it is, and apply formally the Kac-Rice formula to deduce

$$\begin{gathered}\mathbb{E}\big[\,Z[B]^2\,\big]-\mathbb{E}\big[\,Z[B]\,\big]\\ =\int_{B_*^2}\mathbb{E}\big[\,\big|\det\big(\,F'(v_1)\cdot F'(v_2)\,\big)\,\big|\,\big\|\,F(v_1)=F(v_2)=0\,\big]p_{F(v_1)\oplus F(v_2)}(0)dv_1dv_2.\end{gathered}\tag{2.4.1}$$

Recall that

$$p_{F(v_1)\oplus F(v_2)}(0)=\frac{1}{\sqrt{\det\big(\,2\pi\operatorname{Var}\big[\,F(v_1)\oplus F(v_2)\,\big]\,\big)}}$$

We see that as (v_1,v_2) approaches (v,v), the Gaussian vector $F(v_1)\oplus F(v_2)$ approaches the *degenerate* gaussian vector $F(v)\oplus F(v)$. Hence the variance of $F(v_1)\oplus F(v_2)$ degenerates as (v_1,v_2) approaches the diagonal so the term $p_{F(v_1)\oplus F(v_2)}(0)$ is guaranteed to explode near the diagonal. This raises the issue of finiteness of the integral in (2.4.1).

Now that we are guaranteed a headache, let us recall that we still do not know wether the Gaussian vector $F(v_1)\oplus F(v_2)$ is nondegenerate if $v_1\neq v_2$. Fortunately, under additional assumptions of F this will be the case.

In the next, warm-up, subsection we show that, under reasonable assumptions, the variance is finite. The key idea in the proof is the gauge invariance of the Kac-Rice density.

More precisely suppose that

$$\Psi: M\to E$$

is an ample random Gaussian section of the real vector bundle $E\to M$, where M is a smooth manifold. Suppose thatnthe dimension of M is equal to the rank of E. Then, for any gauge transformation (or linear automorphism) $g:E\to E$, the random Gaussian section $g(\Psi)$ is also ample and we have a tautological equality of random measures

$$Z\big[\,-,\Psi\,\big]=Z\big[\,-,g(\Psi)\,\big].$$

For illustration purposes suppose that M is the plane, $M=\mathbb{R}^2$, and E is the trivial rank 2-bundle. Then the section $\Phi(x)=|\boldsymbol{x}|^2\Psi(\boldsymbol{x})$ over the punctured plane is gauge equivalent to Ψ, but the Kac-Rice formula suggests that the Kac-Rice density of Φ might blow-up at the origin since

$$p_{\Phi(x)}(0)=p_{|\boldsymbol{x}|^2\Psi(x)}(0)=|\boldsymbol{x}|^{-4}p_{\Psi(x)}(0).$$

On the other hand,

$$\rho_\Psi(\boldsymbol{x})=\rho_\Phi(\boldsymbol{x}),\;\;\boldsymbol{x}\neq 0$$

since $Z\big[\,S,\Psi\,\big]=Z\big[\,-,g(\Psi)\,\big]=Z\big[\,-,\Phi\,\big]$, for any Borel subset $S\subset\mathbb{R}^2\setminus 0$. The gauge transformation $g(\boldsymbol{x})=|\boldsymbol{x}|^2\mathbb{1}_{\mathbb{R}^2}$ desingularizes Φ in the sense that $\Phi=g\Psi$ and ψ is much better behaved section.

This argument can be slightly generalized. Suppose that we are given two real vector bundles $E_0,E_1\to M$ If $\Psi_0: M\to E_0$ is a Gaussian random section of the real vector bundle

[5]This is sometimes referred to as the Pohamer symbol

$E_0 \to M$, and $T : E_0 \to E_1$ is a bundle isomorphism, then the Gaussian random section $T\Psi_0$ has the same zero set as Ψ_0. However, with a bit of luck, the renormalized section $T\Psi_0$ may be better behaved and free of degenerations of the type mentioned above.

2.4.2. Variance estimates. It has been known for some time that under certain conditions the number of zeros in a box of a Gaussian field F has finite variance, [**3, 16, 55, 66**]. In this warm-up subsection we use the ideas in the above references to obtain such estimates for the variance in terms of the covariance kernel. Here an in the sequel

Suppose that $\boldsymbol{U}$ and $\boldsymbol{V}$ are finite dimensional real Euclidean spaces of the same dimension m and $\mathscr{V} \subset \boldsymbol{V}$ is an open set. If $f : \mathscr{V} \to \boldsymbol{U}$ is a C^k-map, we denote by $f^{(k)}(v)$ its k-th differential at $v \in \mathscr{V}$. We view $f^{(k)}(v)$ as an element of $\mathbf{Sym}^k(\boldsymbol{V}, \boldsymbol{U})$, the space of symmetric k-linear maps $\boldsymbol{V}^k \to \boldsymbol{U}$.

Let $F : \mathscr{V} \to \boldsymbol{U}$ be a Gaussian random field whose covariance kernel $\mathcal{K}_F$ is C^6. In particular, this implies that F is a.s. C^2.

For any box $B \subset \mathscr{V}$ we denote by Z_B the number of zeros of F in B, i.e., $Z_B = Z[B, F]$. Let $\mathscr{V}_*^2 := \mathscr{V}^2 \setminus \Delta$, where Δ is the diagonal. Define B_*^2 in a similar fashion. Consider the random field

$$\widehat{F} =: \mathscr{V}_*^2 \to \boldsymbol{U} \oplus \boldsymbol{U}, \;\; \hat{F}(v_0, v_1) = F(v_0) \oplus F(v_1).$$

Note that

$$Z[\widehat{F}, B_*^2] = Z_B\big(Z_B - 1\big).$$

Suppose that $F|_B$ is 2-ample, i.e., for any $\underline{v} = (v_0, v_1) \in B_*^2$ the Gaussian vector $F(v_0) \oplus F(v_1)$ is nondegenerate. We deduce from the local Kac-Rice formula (**KR**) that $E\big[\, Z_B \,\big] < \infty$, and

$$\mathbb{E}\big[\, Z_B\big(\, Z_B - 1\,\big)\,\big] = \int_{B_*^2} \rho_G^{(2)}(v_0, v_1) dv_0 dv_1,$$

where $\rho_F^{(2)}$ is the Kac-Rice density

$$\rho_F^{(2)}(v_0, v_1) := \mathbb{E}\big[\, |\det F'(v_0) \det F'(v_1)| \,\big|\, F(v_0) = F(v_1) = 0\,\big] p_{\widehat{F}(v_0, v_1)}(0). \tag{2.4.2}$$

Note that

$$p_{\widehat{F}(v_0,v_1)}(0) = \frac{1}{\sqrt{\det\big(\, 2\pi \operatorname{Var}[F(v_0) \oplus F(v_1)]\,\big)}},$$

so $p_{\widehat{F}(v_0,v_1)}(0)$ explodes as (v_0, v_1) approaches the diagonal since $F(v) \oplus F(v)$ is degenerate for any $v \in \mathscr{V}$. Thus the function $\rho_F^{(2)}(v_0, v_1)$ might have a non-integrable singularity along the diagonal so $E\big[\, Z_B^2\,\big]$ could be infinite.

We want to show that this is not the case and a bit more. We will use the gauge-change trick outlined in the introduction to his section.

Proposition 2.4.1. *Fix a box $B \subset \mathscr{V}$ and $r < \operatorname{dist}(B, \mathscr{V})$. Denote by $S = S(r, B)$ the compact set set*

$$S = \big\{\, v \in \mathscr{V};\;\; \operatorname{dist}(v, B) \leq r \,\big\}.$$

Suppose that $F\big|_B$ is C^2, 2-ample and J_1-ample, i.e., for any $\boldsymbol{v} \in B$ the Gaussian vector $\big(\, F(v), F'(v)\,\big)$ is nondegenerate. Define

$$w_F : B_*^2 \to \mathbb{R}, \;\; w_F(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|^{m-2} \rho_F^{(2)}(\boldsymbol{x}, \boldsymbol{y}).$$

There exists a constant $C(m, \mathrm{vol}[B], r) > 0$*, that depends only on* m*,* $\mathrm{vol}[B]$ *and* r *such that*

$$\sup_{\boldsymbol{p}\in B_*^2} |w_F(\boldsymbol{p})\,| \leq C(m, \mathrm{vol}[B], r)\|\mathcal{K}_F\|^{3m-1/2}_{C^6(S\times S)}. \tag{2.4.3}$$

In particular $\mathrm{Var}\big[\,Z_B\,\big] < \infty$*.*

Proof. I will use a modification of the arguments in [**16**, Sec. 4.2]. For any $v_0, v_1 \in B$, $v_0 \neq v_1$, the Gaussian vector $\hat{F}(v_0, v_1) = F(v_0) \oplus F(v_1)$ is nondegenerate. We denote by $p_{F(v_0),F(v_1)}$ the probability density of $\hat{F}(v_0, v_1)$.

We set

$$r(\underline{v}) := \|v_1 - v_0\|,\ \ \Xi(\underline{v}) := \frac{1}{r(\underline{v})}\big(\,F(v_1) - F(v_0)\,\big).$$

Note that

$$\hat{F}(\underline{v}) = 0 \Longleftrightarrow F(v_0) = \Xi(\underline{v}) = 0.$$

Denote by $A(\underline{v})$ the linear map $\boldsymbol{U}^2 \to \boldsymbol{U}^2$ given by

$$A(\underline{v})\left[\begin{array}{c} u_0 \\ u_1 \end{array}\right] = \left[\begin{array}{c} u_0 \\ u_0 + r(\underline{v})u_1 \end{array}\right] = \left[\begin{array}{cc} \mathbb{1}_{\boldsymbol{U}} & 0 \\ \mathbb{1}_{\boldsymbol{U}} & r(\underline{v})\mathbb{1}_{\boldsymbol{U}} \end{array}\right]\cdot\left[\begin{array}{c} u_0 \\ u_1 \end{array}\right]. \tag{2.4.4}$$

Thus

$$\left[\begin{array}{c} F(v_0) \\ F(v_1) \end{array}\right] = A(\underline{v})\left[\begin{array}{c} F(v_0) \\ \Xi(\underline{v}) \end{array}\right].$$

The gauge transformation $A(\underline{v})$ desingularizes $\hat{F}$. Denote by $Z(\underline{v})$ the Gaussian vector $(F(v_0), \Xi(\underline{v}))$.

The Gaussian regression formula implies that

$$\begin{aligned} &\mathbb{E}\big[\,|\det F'(v_0)\det F'(v_1)|\;\big|\;F(v_0) = F(v_1) = 0\,\big] \\ &\qquad = \mathbb{E}\big[\,|\det F'(v_0)\det F'(v_1)|\;\big|\;Z(\underline{v}) = 0\,\big]. \end{aligned}$$

Note that

$$p_{F(v_0),F(v_1)} = \frac{1}{\sqrt{\det\big(\,2\pi\,\mathrm{Var}[F(v_0)\oplus F(v_1)]\,\big)}}$$

$$= \frac{1}{|\det A|\sqrt{\det\big(\,2\pi\,\mathrm{Var}[F(v_0)\oplus \Xi(\underline{v}))]\big)}} = r(\underline{v})^{-m}p_{F(v_0)\oplus\Xi(\underline{v})}(0).$$

We deduce that for any $\underline{u} \in B_*^2$ we have

$$\rho_F^{(2)}(\underline{v}) := r(\underline{v})^{-m}\mathbb{E}\big[\,|\det F'(v_0)\det F'(v_1)|\;\big|\;Z(\underline{v}) = 0\,\big]p_{F(v_0)\oplus\Xi(\underline{v})}(0). \tag{2.4.5}$$

Lemma 2.4.2. *There exists a constant* $C = C(m, \mathrm{vol}[B], r) > 0$ *depending only on* m *and* $\mathrm{vol}[B]$ *and* $r < \mathrm{dist}(B, \partial\mathscr{V})$ *such that, for* $i = 0, 1$*, and any* $\underline{v} \in B_*^2$

$$\Big|\,\mathbb{E}\big[\,|\det F(v_i)|^2\big|\,Z(\underline{v}) = 0\,\big]\,\Big| \leq C(m, \mathrm{vol}[B], r)\|\mathcal{K}_F\|^{m+2}_{C^6(S\times S)}r(\underline{v})^2.$$

Proof. It suffices to consider only the case $i = 0$ since

$$F(v_0) = \Xi(\underline{v}) = 0 \Longleftrightarrow F(v_1) = \Xi(\underline{v}) = 0.$$

Set

$$\nu = \nu(\underline{v}) := \frac{1}{r(\underline{v})}\big(\,v_1 - v_0\,\big),\ \ Z = Z(\underline{v}).$$

Let $f(t) = F(v_0 + t\nu)$. Since $F(v)$ is a.s. C^2 we deduce from the first order Taylor approximation with integral remainder that

$$F(v_1) - F(v_0) = f(r) - f(0) = rf'(0) + \int_0^r f''(t)(r-t)dt = \partial_\nu F(v_0) + \underbrace{\int_0^r f''(t)(r-t)dt}_{=:W}.$$

Hence

$$r\partial_\nu F(v_0) = F(v_0) - F(v_1) - W$$

Hence, for any $p \geq 1$ we have

$$\mathbb{E}\big[\,|r\partial_\nu F(v_0)|^p \,\big|\, Z=0\,\big] = \mathbb{E}\big[\,|F(v_0) - F(v_1) - W|^p \big|\, Z=0\,\big] = \mathbb{E}\big[\,|W|^p \,\big|\, Z=0\,\big].$$

The random variable W is a centered $\boldsymbol{U}$-valued Gaussian vector. We deduce that for any $p \geq 1$ we have

$$\big|\,\mathbb{E}\big[\,|\partial_\nu F(v_0)|^p \big|\, Z=0\,\big]\,\big| = \frac{1}{r^p}\mathbb{E}\big[\,|W|^p \,\big|\, Z=0\,\big]^p.$$

Note that

$$|W| \leq \int_0^r \|f''(t)\|_{\boldsymbol{U}}(r-t)dt \leq \frac{r^2}{2}\|F\|_{C^2(B)}.$$

We deduce that

$$\big\|\ \mathrm{Var}\,\big[\,W\,\big]\,\big\|_{\mathrm{op}} \leq \frac{r^4}{4}\mathbb{E}\big[\,\|F\|^2_{C^2(B)}\,\big].$$

Using Corollary 1.1.30 we deduce that

$$\mathbb{E}\big[\,|W|^p \,\big|\, Z=0\,\big] \leq C(m,p)r^{2p}\mathbb{E}\big[\,\|F\|^2_{C^2(B)}\,\big]^{p/2},$$

where $C(m,p)$ is a universal constant that depends only on the dimension m and on p. We will continue to denote by the same symbol $C(m,p)$ various positive constants that depend only on m and p. We deduce

$$\big|\,\mathbb{E}\big[\,\partial_\nu F(v_0) \big|\, Z=0\,\big]\,\big|^p \leq C(m,p)r^p\mathbb{E}\big[\,\|F\|^2_{C^2(B)}\,\big]^{p/2}. \tag{2.4.6}$$

Extend ν to an orthonormal basis $\{\nu = \boldsymbol{e}_1, \boldsymbol{e}_2, \dots \boldsymbol{e}_m\}$ of $\boldsymbol{V}$. Using Hadamard's inequality [**75**, Cor. 7.8.2] we deduce

$$\begin{aligned}\big|\,\det F'(v_0)\,\big| &= \big|\,\det\big(\,\partial_{\boldsymbol{e}_1}F(v_0), \partial_{\boldsymbol{e}_2}F(v_0), \dots, \partial_{\boldsymbol{e}_m}F(v_0)\,\big)\,\big| \\ &\leq \big|\,\partial_{\boldsymbol{e}_1}F(v_0)\,\big|\prod_{k=2}^m\big|\,\partial_{\boldsymbol{e}_k}F(v_0)\,\big|.\end{aligned}$$

Using Hölder's inequality we deduce

$$\mathbb{E}\big[\,\big|\,\det F'(v_0)\,\big|^2\,\big|Z=0\ \big] \leq \prod_{k=1}^m\mathbb{E}\big[\,\big|\,\partial_{\boldsymbol{e}_k}F(v_0)\,\big|^{2m}\,\big|\,\big|\,Z=0\,\big]\,\big]^{\frac{1}{m}}.$$

For $k = 2, \dots, m$ we have

$$\mathrm{Var}\,\big[\,\partial_{\boldsymbol{e}_k}F(v_0)\,\big|\,Z=0\,\big] \leq \mathrm{Var}\,\big[\,\partial_{\boldsymbol{e}_k}F(v_0)\,\big]$$

and

$$\big\|\ \mathrm{Var}\,\big[\,\partial_{\boldsymbol{e}_k}F(v_0)\,\big]\,\big\|_{\mathrm{op}} \leq C(m)\|\mathcal{K}_F\|_{C^2(B\times B)}.$$

Using again Corollary 1.1.30 we deduce that for $k = 2, \dots, m$ we have

$$\mathbb{E}\big[\,\big|\,\partial_{\boldsymbol{e}_k}F(v_0)\,\big|^{2m}\,\big|Z=0\,\big]^{\frac{1}{m}} \leq C(m)\|\mathcal{K}_F\|_{C^2(B\times B)}.$$

Using (2.4.6), we deduce that

$$\mathbb{E}\big[\,\big|\det F'(v_0)\big|^2 \,\big|\, Z=0\,\big] \le C(m)r^2\mathbb{E}\big[\,\|F\|^2_{C^2(B)}\,\big]\|\mathcal{K}_F\|^{m-1}_{C^2(B\times B)}.$$

Invoking (1.2.4) we conclude that

$$\mathbb{E}\big[\,\|F\|^2_{C^2(B)}\,\big] \le C(m,\operatorname{vol}[B],r)\|K\|^3_{C^6(S\times S)}.$$

This completes the proof of Lemma 2.4.2. □

Lemma 2.4.2 implies

$$\begin{gathered}\mathbb{E}\big[\,|\det F'(v_0)\det F'(v_1)|\,\big|\,Z(\underline{v})=0\,\big]\\ \le \mathbb{E}\big[\,\big|\det F'(v_0)\big|^2\,\big|Z(\underline{v})=0\,\big]\,\big|^{1/2}\mathbb{E}\big[\,\big|\det F'(v_1)\big|^2\,\big|\,Z(\underline{v})=0\,\big]\,\big|^{1/2}\\ \le C(m,\operatorname{vol}[B],r)\|K\|^{m-1/2}_{C^6(S\times S)}r(\underline{v})^{-2}.\end{gathered}$$

Hence

$$\rho_F^{(2)}(\underline{v}) \le C(m)\|K\|^{m-1/2}_{C^6(S\times S)}r(\underline{v})^{2-m}\sup_{\underline{v}} p_{F(v_0)\oplus\Xi(\underline{v})}(0). \tag{2.4.7}$$

Moreover

$$\sup_{\underline{v}} p_{F(v_0)\oplus\Xi(\underline{v})}(0) \le C(m)\|K\|^{2m}_{C^3(B\times B)} \le C(m)\|K\|^{2m}_{C^6(S\times S)}.$$

This completes the proof of Proposition 2.4.1. □

We can extract from the above proof a more precise result. For any box B in a Euclidean space $\boldsymbol{V}$ we set

$$\mathfrak{q}(B) := \int_{B_*^2} r(\underline{v})^{2-m}dv_0dv_1.$$

Note that $\mathfrak{q}(B)$ is a translation invariant and for any $t>0$, $\mathfrak{q}(tB) = t^{m+2}\mathfrak{q}(B)$. In particular, if B is is the cube $B_c = [0,c]^m$, then

$$\mathfrak{q}(B_c) = \mathfrak{q}(B_1)c^{m+2} = C(m)\mathfrak{q}(B_1)\operatorname{vol}\big[\,B_c\,\big]^{\frac{m+2}{m}}.$$

Corollary 2.4.3. *Let $\mathcal{V}$ be an open subset of $\boldsymbol{V}$. For each $r>0$ there exists a function*

$$\mathfrak{F} : (0,\infty) \to (0,\infty)$$

with the following property: for any $m_0 > 0$, any box $B \subset \mathscr{V}$ and any Gaussian field $F : \Omega \times \mathcal{V} \to \boldsymbol{U}$ such that

- $\operatorname{dist}(B, \partial\mathscr{V}) < r$,
- *the covariance kernel $\mathcal{K}_F$ is C^6,*
- *the restriction of F to B is 2-ample,*
- *and $\big\|\,\mathcal{K}\,\big\|_{C^6(S\times S)} < m_0$*

we have

$$\big\|\,\rho^{(2)}_{KR}\,\big\|_{L^1(B\times B)} < \mathfrak{F}_r(m_0)\mathfrak{q}(B).$$

□

Remark 2.4.4. One can show that if F is a.s. C^3, then the function w_F in Proposition 2.4.1 admits an extension to a continuous function on the radial blow-up of B^2 along the diagonal. □

2.4.3. An analytic digression: Kergin interpolation. The one-dimensional case of this technique goes back to Newton.

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuous function and $x_1, \dots, x_p$ are *distinct points* on the real axis. We define inductively the *divided differences* $f[x_1], f[x_1, x_2], \dots, f[x_1, \dots, x_p]$ by setting

$$f[x] = f(x), \;\; \forall x \in \mathbb{R},$$

$$f[x_1, x_2] = \frac{f[x_1] - f[x_2]}{x_1 - x_2} = f[x_2, x_1],$$

$$f[x_1, x_2, x_3] = \frac{f[x_1, x_2] - f[x_2, x_3]}{x_1 - x_3} = \frac{f[x_3, x_2] - f[x_1, x_2]}{x_3 - x_1},$$

$$f[x_1, x_2, \dots, x_k, x_{k+1}] = \frac{f[x_1, \dots, x_k] - f[x_2, \dots, x_{k+1}]}{x_1 - x_{k+1}} \dots$$

For simplicity we will write $\underline{x} = (x_1, \dots, x_p)$ and $f[\underline{x}] = f[x_1, \dots, x_p]$. For distinct $x_1, \dots, x_n$ we have the following more explicit description (see [**102**, Sec. 1.3])

$$f[x_1, \dots, x_n] = \sum_{j=1}^{n} \frac{f(x_j)}{\prod_{k \neq j}(x_j - x_k)}.$$

If $f \in C^p$, then we have an alternate integral representation of $f[x_0, \dots, x_p]$ called *Hermite-Genocchi formula*

$$f[x_0, x_1, \dots, x_p] = \int_0^1 ds_1 \int_0^{s_2} ds_3 \cdots \int_0^{s_{p-1}} f^{(p)}\big(\, y(\underline{s})\,\big) ds_p, \tag{2.4.8}$$

where

$$y(\underline{s}) = y(s_1, \dots, s_p) = (1 - s_1)x_0 + (s_1 - s_2)x_1 + \cdots + (s_{p-1} - s_p)x_p,$$
$$1 \geq s_1 \geq \cdots \geq s_p \geq 0.$$

We refer to [**24**, Thm. 1.9] or [**102**, Sec. 16] for a proof. Note that this formula assumes that f is p-times differentiable. We can rephrase (2.4.8) in more revealing terms as follows.

Consider the simplex

$$\Delta_p = \Big\{ \,(t_0, t_1, \dots, t_p) \in [0, 1]^{p+1}; \;\; \sum_{k=0}^{p} t_k = 1 \,\Big\}.$$

The symmetric group $\mathfrak{S}_{p+1}$ acts on Δ_p by permuting the variables $t_0, \dots, t_p$. Moreover, Δ_p is equipped with an Euclidean volume element $\mathrm{vol}\left[\,-\,\right]$ induced by the Euclidean inner product. The volume element $\mathrm{vol}\left[\,-\,\right]$ is invariant with respect to the action of $\mathfrak{S}_{p+1}$

We view Δ_p as graph of the function $t_0 = 1 - (t_1 + \cdots + t_p)$. We can use $(t_1, \dots t_p)$ as local coordinates and we deduce

$$\mathrm{vol}\left[\, dt_1 \cdots dt_p \,\right] = \sqrt{1 + |\nabla t_0|^2} dt_1 \dots dt_p = \sqrt{p+1} dt_1 \dots dt_p.$$

We have

$$\mathrm{vol}\left[\, \Delta_p \,\right] = \sqrt{p+1} \underbrace{\int_{\substack{t_1, \dots, t_p \geq 0 \\ t_1 + \cdots + t_p \leq 1}} dt_1 \cdots dt_p}_{=\frac{1}{p!}} = \frac{\sqrt{p+1}}{p!}.$$

Let

$$\mu_p\big[\,dt\,\big] := \frac{1}{\sqrt{p+1}}\,\mathrm{vol}\big[\,dt\,\big] = dt_1\cdots dt_p,$$

so that $\mu_p\big[\,\Delta_p\,\big] = \frac{1}{p!}$. Given $\underline{x} = (x_0, x_1, \dots, x_p) \in \mathbb{R}^{p+1}$ we define

$$\sigma_{\underline{x}} : \Delta_p \to \mathbb{R},\ \ \sigma_{\underline{x}}(t) := \sum_{k=0}^{p} t_k x_k.$$

If we make the linear change in variables $s_k = t_{p-k+1} + \cdots + t_p$, $1 \leq k \leq p$, $t_0 = 1 - s_1$, then for any continuous function $u : \mathbb{R} \to \mathbb{R}$ we have

$$\begin{aligned}\int_{\Delta_p} u\big(\,\sigma_{\underline{x}}(t)\,\big)\mu_p\big[\,dt\,\big] &= \int_{\substack{t_1,\dots,t_p\geq 0\\ t_1+\cdots+t_p\leq 1}} u\big(\,\sigma_{\underline{x}}(t)\,\big)dt_1\cdots dt_p\\ &= \int_0^1 ds_1\int_0^{s_2} ds_3\cdots\int_0^{s_{p-1}} u\big(\,y(\underline{s})\,\big)ds_p,\end{aligned}$$

$$y(\underline{s}) = (1-s_1)x_0 + (s_1-s_2)x_1 + \cdots + (s_{p-1}-s_p)x_p$$

Then (2.4.8) can be rewritten as

$$f[\underline{x}] = \int_{\Delta_p} f^{(p)}\big(\,\sigma_{\underline{x}}(t)\,\big)\mu_p\big[\,dt\,\big]. \tag{2.4.9}$$

The measure μ_p is invariant under the action of the symmetric group $\mathfrak{S}_{p+1}$ so the right-hand-side of the above equality is symmetric in the variables $x_0, \dots, x_p$. It also depends continuously on them *and it is well defined even if some of them coincide*!

This allows us to define $f[x_0, x_1, \dots, x_p]$ even if the numbers $x_0, \dots, x_p$ are not pairwise distinct provided that $f \in C^p$. For example,

$$f[x_1, x_1] = \lim_{x_2\to x_1} f[x_1, x_2] = f'(x_1),$$

$$f[x_1, x_1, x_2] = \frac{f[x_2, x_1] - f'(x_1)}{x_2 - x_1}$$

More generally, if the function $f(x)$ is C^k, then the function $g(x) = f[x, x_2]$ is C^{k-1} and

$$f[x_1, x_2, x_3] = g[x_1, x_3] = \frac{f[x_1, x_2] - f[x_3, x_2]}{x_1 - x_3}.$$

In general, for distinct $x, x_1, \dots, x_p$, we have the equality (see [**102**, Sec. 1.1])

$$\begin{aligned}f(x) = \underbrace{f(x_1) + \sum_{j=1}^{p-1}(x-x_1)\cdots(x-x_j)f[x_1,\cdots,x_{j+1}]}_{=:\boldsymbol{P}_{x_1,\dots,x_p}f(x)}\\ +(x-x_1)\cdots(x-x_p)f[x, x_1, \dots, x_p].\end{aligned} \tag{2.4.10}$$

The term $\boldsymbol{P}_{x_1,\dots,x_p}f(x)$ is a polynomial of degree $\leq (p-1)$ in x and the above formula is called *Newton's interpolation formula.* The above equality shows that

$$\boldsymbol{P}_{x_1,\dots,x_p}f(x_i) = f(x_i),\ \ \forall i = 1, \dots, p.$$

As mentioned earlier, the divided difference $f[x_1, \dots, x_p]$ is well defined *even if the numbers* $x_1, \dots, x_p$ *are not pairwise distinct* and thus (2.4.10) holds for any $x, x_1, \dots, x_p \in \mathbb{R}$, provided that $f \in C^p$. Note that if $x_1 = \cdots = x_m$, then (2.4.10) implies that

$$\partial_x^k \boldsymbol{P}_{x_1,\dots,x_m} f(x_1) = \frac{1}{k!}\partial_x^k f(x_1) \;\; \forall 0 \leq k < m.$$

If we set

$$[x_0]_m := \underbrace{x_0, \dots x_0}_{m},$$

then

$$\boldsymbol{P}_{[x_0]_m}(x) = \sum_{j=1}^m \frac{1}{(j-1)!} f^{(j-1)}(x - x_0)^{j-1}.$$

is the degree $m-1$ Taylor polynomial of f at x_0.

Let us observe that for f continuous and injective

$$\underline{x} : \mathbb{I}_p := \{\, 1, \dots, p \,\} \to \mathbb{R}$$

the polynomial $Q = \boldsymbol{P}_{\underline{x}} f$ is the Lagrange interpolation polynomial, i.e., the unique polynomial Q of degree $\leq p-1$ such that

$$Q(x_i) = f(x_i), \;\; \forall i = 1, \dots, p.$$

This proves that $P_{\underline{x}}$ is a linear projector, i.e.,

$$\boldsymbol{P}_{\underline{x}}^2 f = \boldsymbol{P}_{\underline{x}} f \in \mathbb{R}[\, x \,], \;\; \forall f \in C(\mathbb{R}),$$

and that $\boldsymbol{P}_{\underline{x}}$ is invariant under the action of $\mathfrak{S}_p$ on $\mathbb{R}^p$. Moreover, for any $I \subset \mathbb{I}_p$ we have

$$\boldsymbol{P}_{\underline{x}} f(\underline{x}_I) = f(\underline{x}_I).$$

The continuous dependence $\underline{x} \to \boldsymbol{P}_{\underline{x}}$ shows that, for any $\underline{x} \in \mathbb{R}^p$ and any $I \subset \mathbb{I}_p$, the map $\mathbb{P}_{\underline{x}}$ is a symmetric linear projector of $C^{p-1}(\mathbb{R})$, i.e., for any permutation $\varphi \in \mathfrak{S}_p$

$$\boldsymbol{P}_{\underline{x}}^2 f = \boldsymbol{P}_{\underline{x}} f = \boldsymbol{P}_{\underline{x}\cdot\varphi} f, \;\; \forall f \in C^{p-1}(\mathbb{R}), \tag{2.4.11}$$

and

$$\boldsymbol{P}_{\underline{x}} = \boldsymbol{P}_{\underline{x}_I}. \tag{2.4.12}$$

Formula (2.4.9) is the basis of the higher dimensional generalization of the above classical facts, [**82, 101**].

Fix an m-dimensional Euclidean space $\boldsymbol{V}$ and $\mathscr{V} \subset \boldsymbol{V}$ an open *convex* subset. Given a function $f \in C^p(\mathscr{V})$ and $1 \leq k \leq p$, the k-th differential of f at $v \in \mathscr{V}$, denoted by $D^k f(v)$, is a symmetric k-linear form on $\boldsymbol{V}$,

$$D^k f(v) \in \mathbf{Sym}_k(\boldsymbol{V}).$$

Given $\underline{v} = (v_0, v_1, \dots, v_k) \in \mathscr{V}^{k+1}$ we define

$$\sigma_{\underline{v}} = \sigma_{\underline{v}}^k : \Delta_k \to \mathscr{V}, \;\; \sigma_{\underline{v}}(t) := \sum_{i=0}^k t_i v_i,$$

and

$$f[\underline{v}] := \int_{\Delta_k} D^k f\big(\, \sigma_{\underline{v}}(t)\,\big)\mu_k\big[\, dt \,\big] \in \mathbf{Sym}_k(\boldsymbol{V}).$$

Given $v_0, v_1, \dots, v_p \in \mathscr{V}$ we define the *Kergin interpolator* of f to be the polynomial of degree $\leq p$ in u,

$$\boldsymbol{P}_{v_0,v_1,\dots,v_p} f(v) = f(v_0) + \sum_{k=1}^{p} \underbrace{f[v_0,\dots,v_k]}_{\in \mathbf{Sym}_k(\boldsymbol{V})} \big(v - v_0, \dots, v - v_{k-1} \big). \tag{2.4.13}$$

For example, when $p = 1$ we have

$$\boldsymbol{P}_{v_0,v_1} f(v) = f(v_0) + f[v_0, v_1](v - v_0) = f(v_0) + \int_0^1 \partial_{(v-v_0)} f\big((1-t)v_0 + tv_1 \big) dt, \tag{2.4.14}$$

where ∂_w denotes the directional derivative in the direction w.

Suppose that f is a *ridge function*, i.e., there exists a C^p-function $g : \mathbb{R} \to \mathbb{R}$ and a linear form $\xi \in \boldsymbol{V}^*$ such that $f(v) = g\big(\xi(v)\big)$. Informally, a ridge function depends on a single linear coordinate. Then

$$f\big[\, v_0, \dots, v_k \,\big] = g\big[\, \xi_0, \dots, \xi_k \,\big], \;\; \xi_k = \xi(v_k), \;\; 0 \leq k \leq p.$$

In particular,

$$\boldsymbol{P}_{v_0,v_1,\dots,v_p} f(v) = \boldsymbol{P}_{\xi_0,\dots,\xi_p} g(x), \;\; x = \xi(v)$$

Thus

$$\boldsymbol{P}_{v_0,\dots,v_p} f(v_k) = f(v_k), \;\; \forall 0 \leq k \leq p, \tag{2.4.15}$$

for any function f that is a linear combination of ridge functions. The linear span of ridge functions contains the space of polynomials (see [**24**, Lemma 9.11]) which is dense in $C^p(\mathscr{V}, \mathbb{R})$, so (2.4.15) holds for any $f \in C^p(\mathscr{V})$.

A similar argument shows that $\boldsymbol{P}_{v_0,\dots,v_p} f$ is symmetric in the variables $v_0, v_1, \dots, v_p$.

Given $q \leq p$ and $\underline{v} = (v_0, v_1, \dots, v_p) \in \mathscr{V}^{p+1}$, we set $[\underline{v}]_q := (v_0, \dots, v_q)$. We have

$$\boldsymbol{P}_{[\underline{v}]_q} \boldsymbol{P}_{\underline{v}} = \boldsymbol{P}_{[\underline{v}]_q}. \tag{2.4.16}$$

Indeed, this is true when $d = 1$ and thus it is true for arbitrary d and f a ridge function. The conclusion follows by linearity and density. In particular, when $q = p$ the above equality shows that $\boldsymbol{P}_{\underline{v}}$ is a projector. For this reason we will also refer to $\boldsymbol{P}_{\underline{v}}$ as *Kergin projector*.

Let $p \geq 1$. We denote by $\mathrm{Poly}_p\big[\, \boldsymbol{V} \,\big]$ the vector space of polynomials maps $\boldsymbol{V} \to \mathbb{R}$ of degree $\leq p$. Define

$$m_i : \mathscr{V}^{p+1} \to \mathbb{N}, \;\; m_i(v_0, v_1, \dots, m_p) = \#\big\{\, k; \;\; u_k = u_i \,\big\}.$$

We refer to $m_i(\underline{v})$ the multiplicity of v_i in $\underline{v} = (v_0, \dots, v_p)$, i.e., the number of terms of the sequence of points $v_0, \dots, v_p$ equal to v_i. We have the following result, [**82, 101**].

Theorem 2.4.5. *Let $\underline{v} \in \mathscr{V}^{p+1}$. The map*

$$\boldsymbol{P}_{\underline{v}} : C^p(\mathscr{V}) \to \mathrm{Poly}_p\big[\, \boldsymbol{V} \,\big] \subset C^p(\mathscr{V}), \;\; f \mapsto \boldsymbol{P}_{\underline{v}} f,$$

is a linear continuous projector, i.e., $\boldsymbol{P}_{\underline{v}}^2 = \boldsymbol{P}_{\underline{v}}$. It depends continuously on $\underline{u}$. Moreover, for any $i = 0, 1, \dots, p$ and any multi-index $\alpha \in \mathbb{N}_0^d$ such that $|\alpha| < m_i(p)$ we have

$$\partial^\alpha \boldsymbol{P}_{\underline{v}} f(v_i) = \partial^\alpha f(v_i). \tag{2.4.17}$$

□

The Kergin interpolator extends in an obvious way to maps $F := C^p(\mathscr{V}, \boldsymbol{U})$, where $\boldsymbol{U}$ is a Euclidean space of dimension n. We will denote by $\operatorname{Poly}_p\big[\,\boldsymbol{V},\boldsymbol{U}\,\big]$ the space of polynomial maps $\boldsymbol{V}\to\boldsymbol{U}$ of degree $\leq p$. More precisely

$$P\in \operatorname{Poly}_p\big[\,\boldsymbol{V},\boldsymbol{U}\,\big] \Longleftrightarrow \forall\xi\in\boldsymbol{U}^*,\;\; \xi(P)\in \operatorname{Poly}_p\big[\,\boldsymbol{V}\,\big].$$

For any $\underline{v}\in\mathscr{V}^{p+1}$ the interpolator $\boldsymbol{P}_{\underline{u}}G$ is the unique polynomial map $\mathscr{V}\to\boldsymbol{U}$ of degree $\leq p$ such that, for any linear functional $\xi\in\boldsymbol{U}^*$ we have

$$\xi\big(\,\boldsymbol{P}_{\underline{u}}G\,\big)=\boldsymbol{P}_{\underline{u}}\xi(G).$$

More explicitly, using Euclidean coordinates $(v^1,\dots,v^m)$ on $\boldsymbol{V}$ and Euclidean coordinates $(u^1,\dots,u^n)$ on $\boldsymbol{U}$ we can view F is an n-tuple of functions

$$G=\left[\begin{array}{c} G^1\\ \vdots\\ G^n\end{array}\right],$$

and then

$$\boldsymbol{P}_{\underline{v}}G:=\left[\begin{array}{c} \boldsymbol{P}_{\underline{v}}G^1\\ \vdots\\ \boldsymbol{P}_{\underline{v}}G^n\end{array}\right].$$

A differential 1-form on $\mathscr{V}$ can viewed as a map $\mathscr{V}\to\boldsymbol{V}^*$, and in particular, we can speak of the Kergin interpolator of a differential 1-form.

We have the following result of Gass and Stecconi [66, Lemma 2.5] stating that the Kergin interpolator of an exact form is also exact.

Lemma 2.4.6. *Let $\underline{u}^*=(v_0^*,v_1^*,\dots,v_p^*)\in\mathscr{V}^{p+1}$ and $f\in C^{p+1}(\mathscr{V})$. Then for any $k=0,1,\dots,p$ and any $i,j\in\{1,\dots,m\}$ we have*

$$\partial_{v^j}\big(\,\boldsymbol{P}_{\underline{v}^*}\partial_{v^i}f\,\big)=\partial_{v^i}\big(\,\boldsymbol{P}_{\underline{v}^*}\partial_{v^j}f\,\big).$$

In other words the polynomial vector field

$$(V_1,\dots,V_m)=\boldsymbol{P}_{\underline{v}^*}\nabla f=\big(\,\boldsymbol{P}_{\underline{u}^*}\partial_{v_1}f,\dots,\boldsymbol{P}_{\underline{v}^*}\partial_{v_m}f\,\big)$$

is a gradient vector field, i.e., there exists a polynomial $h\in\mathbb{R}_{p+1}\big[\,\boldsymbol{V}\,\big]$ such that $\nabla h=\boldsymbol{P}_{\underline{u}^*}\nabla f$.

Proof. We first prove that the lemma is true for ridge functions. By choosing the Euclidean coordinates $(u^1,\dots,u^d)$ carefully this means that $f(u)$ has the form $f(u^1,\dots,u^d)=f(u^1)$. In this case the Lemma is obvious since $\boldsymbol{P}_{\underline{u}}f$ it is a polynomial of degree p in u^1. The general case follows from the density in $C^{p+1}(\mathfrak{U})$ of the linear span of ridge functions. □

2.4.4. Multijets. In this and next subsection we will described the desingularization process devised by Ancona and Letendre and explain how it can be used to provide sufficient conditions that guarantee the finiteness of higher momentums of $Z[B,F]$. It is based on the concept of *multijet* introduced by Ancona and Letendre [4].

In truth, we will present only a special case of their construction that suffices for our purposes. To keep the flow of arguments uninterrupted we will omit the proofs of certain technical results from real algebraic geometry. These proofs use "standard"[6] facts from real

[6] I include Hironaka's resolution of singularities theorem among these "standard" facts.

algebraic geometry. A reader familiar with this subject would have little trouble accepting these results.

As in the previous subsections $\boldsymbol{U}, \boldsymbol{V}$ are real Euclidean spaces of the same dimension m. Fix $k \in \mathbb{N}$, $k \geq 2$. For any $n \in \mathbb{N}$ we set

$$\mathbb{I}_n := \{1, \dots, n\}.$$

For any finite set I we have the space $\boldsymbol{V}^I$ consisting of maps $I \to \boldsymbol{V}$ and a configuration space [7]

$$\mathcal{C}_I(\boldsymbol{V}) \subset \boldsymbol{V}^I$$

consisting of *injective* maps $I \to \boldsymbol{V}$. For $I = \mathbb{I}_k$ we set $\mathcal{C}_k(\boldsymbol{V}) := \mathcal{C}_{\mathbb{I}_k}(\boldsymbol{V})$. We denote by Δ the "fat" diagonal

$$\Delta = \Delta_k = \boldsymbol{V}^k \setminus \mathcal{C}_k(\boldsymbol{V}).$$

Let $\mathscr{P}^k = \mathcal{P}^k(\boldsymbol{V})$ denote the space of polynomial maps $f : \boldsymbol{V} \to \mathbb{R}$ of degree $\leq k-1$. Note that

$$\dim \mathcal{P}^k(\boldsymbol{V}) = \sum_{j=0}^{k-1} \binom{\dim \boldsymbol{V} + j - 1}{j}.$$

We can equip $\mathcal{P}^k$ with the inner product

$$(P, Q) = \int_{\boldsymbol{V}} P(v)Q(v)\Gamma_{\boldsymbol{V}}\big[\,dv\,\big], \;\; \forall P, Q \in \mathcal{P}^k$$

where $\Gamma_{\boldsymbol{V}}$ is the canonical Gaussian measure on the Euclidean space $\boldsymbol{V}$.

Each $\underline{v} \in \mathscr{C}_k(\boldsymbol{V})$ defines a *surjective* map

$$\mathbf{Ev}_{\underline{v}} : \mathscr{P}^k \to \mathbb{R}^k, \;\; f \mapsto \big(\, f(v_1), \dots, f(v_k)\,\big).$$

We denote its kernel by $K_{\underline{v}}$. It is a codimension-k subspace of $\mathscr{P}^k$. We denote by $\mathbf{Gr}_k$ the Grassmannian of codimension-k subspaces of $\mathscr{P}^k$. We thus have a smooth cokernel map map

$$\mathbf{cok} : \mathscr{C}_k(\boldsymbol{V}) \to \mathbf{Gr}_k, \;\; \mathscr{C}_k \ni \underline{v} \mapsto \mathbf{cok}(\underline{v}) = K_{\underline{v}}^{\perp} \in \mathbf{Gr}^k\,.$$

Set

$$L_{\underline{v}} : \mathcal{P} \to \mathbb{R}^k, \;\; L_{\underline{v}} = \big(\mathbf{Ev}_{\underline{v}}\,\mathbf{Ev}_{\underline{v}}^*\big)^{-1/2}\,\mathbf{Ev}_{\underline{v}}\,.$$

As explained in Lemma 1.1.35, the map $L_{\underline{v}}^* : \mathbb{R}^k \to \mathcal{P}^k$ is an isometry whose image is $\mathbf{cok}(\underline{v})$.

Let $\boldsymbol{T}^k \to \mathbf{Gr}_k$ be the tautological vector bundle whose fiber over $S \in \mathbf{Gr}_k$ is S. We denote by $\mathfrak{Proj}_S$ the orthogonal of $\mathcal{P}$ onto S.

Denote by $\underline{\mathbb{R}}^k_{\mathscr{C}^k(\boldsymbol{s})}$ the trivial bundle over $\mathscr{C}_k(\boldsymbol{V})$ with fiber $\mathbb{R}^k$. The maps $L_{\underline{v}}^*$ define vector bundle isomorphism

$$L^* : \underline{\mathbb{R}}^k_{\mathscr{C}^k(\boldsymbol{V})} \to \mathbf{cok}^*\,\boldsymbol{T}^k.$$

Equivalently, we have a commutative diagram

$$\begin{array}{ccc} \underline{\mathbb{R}}^k_{\mathscr{C}^k(\boldsymbol{V})} & \xrightarrow{L^*} & \boldsymbol{T}^k \\ \big\downarrow & & \big\downarrow \\ \mathscr{C}^k(\boldsymbol{V}) & \xrightarrow[\mathbf{cok}]{} & \mathbf{Gr}^k \end{array}$$

[7] Configuration of distinct points in $\boldsymbol{V}$ labelled by I.

where the vertical maps define vector bundles and, for each $\underline{v} \in \mathscr{C}^k(\boldsymbol{V})$, the induced map $L^*_{\underline{v}} : \mathbb{R}^k \to \boldsymbol{T}^k_{\mathbf{cok}(\underline{v})} = \mathbf{cok}(\underline{v}) = K^\perp_{\underline{v}}$ is a linear isometric isomorphism.

We denote by Σ the graph of $\mathbf{cok}$, $\Sigma \subset \mathscr{C}_k(\boldsymbol{V}) \times \mathbf{Gr}^k$. We have a commutative "roof"

$$\begin{array}{ccc} & \Sigma & \\ {\scriptstyle\pi}\swarrow & & \searrow{\scriptstyle\Pi} \\ \mathscr{C}_k(\boldsymbol{V}) & \xrightarrow[\mathbf{cok}]{} & \mathbf{Gr}^k \end{array}$$

where π, Π are the natural projections.

We denote by $\overline{\Sigma}$ the closure of Σ in $\boldsymbol{V}^k \times \mathbf{Gr}^k$. We have a natural projection

$$\pi : \overline{\Sigma} \to \boldsymbol{V}^k,$$

that is algebraic in nature. We can be more precise [**4**, Sec. 5.1].

Proposition 2.4.7. *The following hold.*

(i) *Σ is a smooth real algebraic manifold and the projection $\pi : \Sigma \to \mathscr{C}_k(\boldsymbol{V})$ is a diffeomorphism.*

(ii) *$\overline{\Sigma}$ is a real algebraic variety and the map $\pi : \overline{\Sigma} \to \boldsymbol{V}^k$ is proper and surjective.*

(iii) *The singular locus of $\overline{\Sigma}$ is contained in $\bar{\Delta} : \pi^{-1}(\Delta) = \overline{\Sigma} \setminus \Sigma$, where we recall that $\Delta \in \boldsymbol{V}^k$ denotes the "fat" diagonal.*

□

Invoking Hironaka's (embedded) resolution of singularities theorem one can prove the following result, [**4**, Sec. 5.1] or [**38**, Thm. 6.37].

Theorem 2.4.8. *There exists a smooth manifold W and a proper smooth map*

$$\mathcal{R} : W \to \boldsymbol{V}^k \times \mathbf{Gr}^k$$

with the following properties.

(i) *$\widehat{\Sigma} = \mathcal{R}^{-1}\big(\bar{\Sigma}\big) \subset W$ is smooth.*

(ii) $\dim W = \dim \boldsymbol{V}^k \times \dim \mathbf{Gr}_k$, $\dim \widehat{\Sigma} = \dim \mathscr{C}_k(\boldsymbol{V}) = km = k \dim \boldsymbol{V}$.

(iii) *The set $W^* = W \setminus \mathcal{R}^{-1}\big(\mathcal{C}_k \times \mathbf{Gr}_k\big)$ is open and dense in W and the restriction of $\mathcal{R}$ to W^* is a diffeomorphism onto $\mathcal{C}_k \times \mathbf{Gr}_k$.*

(iv) *The set $\Sigma^* := \mathcal{R}^{-1}\big(\Sigma\big)$ is open and dense in $\hat{\Sigma}$ and the restriction of $\mathcal{R}$ to $\mathcal{R}^{-1}\big(\Sigma\big) \to \Sigma$ is a diffeomorphism onto Σ.*

(v) *The map $\hat{\pi} := \pi \circ \mathcal{R} : \hat{\Sigma} \to \boldsymbol{V}^k$ is smooth and proper. We will refer to the set $\hat{\Delta} := \hat{\pi}^{-1}(\Delta)$ as the* exceptional locus.

□

We set $\hat{\pi} := \pi \circ \mathcal{R}$ and $\hat{\Pi} = \Pi \circ \mathcal{R}$ so that we have a commutative diagram

$$\begin{array}{ccccc} & & \hat{\Sigma} & & \\ & \overset{\hat{\pi}}{\swarrow} & \downarrow \mathcal{R} & \overset{\hat{\Pi}}{\searrow} & \\ \boldsymbol{V}^k & \xleftarrow[\pi]{} & \bar{\Sigma} & \xrightarrow[\Pi]{} & \mathbf{Gr}^k \end{array}$$

The manifold $\hat{\Sigma}$ can be viewed as the graph of a multivalued map

$$\widehat{\mathbf{cok}} : \hat{\Pi} \circ \hat{\pi}^{-1} : \boldsymbol{V}^k \dashrightarrow \mathbf{Gr}^k$$

whose restriction to $\mathscr{C}_k(\boldsymbol{V})$ is the map **cok**.

The pair $(\hat{\Sigma}, \mathcal{R})$ with the above properties is called a resolution of **cok** and is not unique. We fix a resolution and we denote it by $(\widehat{\mathscr{C}}_k, \mathcal{R})$. We set

$$\hat{\pi} := \pi \circ \mathcal{R}, \;\; \widehat{\mathscr{C}}_k^* = \widehat{\mathscr{C}}_k \setminus \hat{\Delta},$$

and we can identify $\widehat{\mathscr{C}}_k^*$ with $\mathscr{C}_k(\boldsymbol{V})$ using the diffeomorphism $\hat{\pi} : \widehat{\mathscr{C}}_k^* \to \mathscr{C}_k$. For any $\underline{v} \in \boldsymbol{V}^p$ we will denote by $\hat{\underline{v}}$ a point in $\hat{\pi}^{-1}(\underline{v}) \in \widehat{\mathscr{C}}_p$. If $\underline{v} \in \mathscr{C}_p(\boldsymbol{V})$, there is only one $\hat{\underline{v}} \in \widehat{\mathscr{C}}_k$ such that $\pi(\hat{\underline{v}}) = \underline{v}$.

Pulling back $\boldsymbol{T}^k$ via $\hat{\Pi}$ we obtain a rank k-vector budle over $\widehat{\mathscr{C}}_k$,

$$\mathscr{M}_k := \hat{\Pi}^*(\boldsymbol{T}^k) \to \widehat{\mathscr{C}}_p.$$

The vector bundle $\mathscr{M}_k$ is the *bundle of k-multijets*. The fiber of $\mathcal{M}_k$ over $\hat{\underline{v}} \in \widehat{\mathscr{C}}_k$ is

$$\mathcal{M}_k(\hat{\underline{v}}) = \mathbf{cok}\left(\hat{\pi}(\hat{\underline{v}})\right) = K^{\perp}_{\hat{\pi}(\hat{\underline{v}})}.$$

To a function $f \in C^k(\boldsymbol{V})$ we can associate a C^1-section of the trivial bundle $\underline{\mathscr{P}^k}_{\widehat{\mathscr{C}}_k}$ (see page iii) namely the family of Kergin projectors

$$\widehat{\mathscr{C}}_k \ni \hat{\underline{v}} \mapsto \boldsymbol{P}_{\hat{\pi}(\hat{\underline{v}})} f \in \mathscr{P}^k.$$

This projects to a C^1 section $\mu_k[f]$ of the bundle of multijets $\mathscr{M}_k$,

$$\mu_k[f](\hat{\underline{v}}) = \mathfrak{Proj}_{\hat{\Pi}(\hat{\underline{v}})} \boldsymbol{P}_{\hat{\pi}(\hat{\underline{v}})} f \in \mathcal{M}_k(\hat{\underline{v}}).$$

Note that for any $\underline{v} \in \mathscr{C}^k(\boldsymbol{V})$ we have

$$\mathbf{Ev}_{\underline{v}}(f) = 0 \Longleftrightarrow \mu_p[f]\big(\,\underline{v}\,\big) = 0.$$

More generally, given a finite dimensional Euclidean space $\boldsymbol{U}$ we have a space $\mathcal{P}^k(\boldsymbol{U}) = \mathcal{P}^k(\boldsymbol{V}, \boldsymbol{U})$ of polynomials maps $\boldsymbol{V} \to \boldsymbol{U}$ of degree $\leq k-1$. For each $\underline{v} \in \boldsymbol{V}^k$ we have a surjection

$$\mathbf{Ev}_{\underline{v}} : \mathcal{P}^k(\boldsymbol{U}) \to \boldsymbol{U}^k,$$

with kernel $K_{\underline{v}}(\boldsymbol{U})$. This is a subspace of $\mathcal{P}^k(\boldsymbol{U})$ of codimension $k \dim \boldsymbol{U}$. We denote by $\mathbf{cok}_{\boldsymbol{U}}(\underline{v})$ its orthogonal complement in $\mathcal{P}^k(U)$. Denote by $\mathbf{Gr}_k(\boldsymbol{U})$ the Grassmannian of subspaces of dimension $k \dim \boldsymbol{U}$ in $\mathcal{P}^k(\boldsymbol{U})$ and by $\boldsymbol{T}^k(\boldsymbol{U}) \to \mathbf{Gr}_k(\boldsymbol{U})$ the tautological vector bundle.

We have a cokernel map

$$\mathbf{cok}_{\boldsymbol{U}} : \mathscr{C}_k \to \mathbf{Gr}_k(\boldsymbol{U})$$

with graph $\Sigma \subset \mathscr{C}_k \times \mathbf{Gr}_k(\boldsymbol{U})$. Fix a resolution $\big(\widehat{\mathscr{C}}_k, \mathcal{R}\big)$ of $\mathbf{cok}_{\boldsymbol{U}}$ as before

We obtain a trivial vector bundle

$$\underline{\mathscr{P}^k(\boldsymbol{U})}_{\widehat{\mathscr{C}}_k} = \big(\, \mathscr{P}^k(\boldsymbol{U}) \times \widehat{\mathscr{C}}_k \to \widehat{\mathscr{C}}_k \,\big),$$

and a bundle of multijets

$$\mathscr{M}_k(\boldsymbol{U}) := \hat{\Pi}^*(\boldsymbol{T}^k(\boldsymbol{U})) \hookrightarrow \underline{\mathscr{P}^k(\boldsymbol{U})}_{\widehat{\mathscr{C}}_k}.$$

The fiber of $\mathscr{M}_k(\boldsymbol{U})$ over $\hat{\underline{v}}$ is

$$\mathscr{M}_k(\boldsymbol{U})(\hat{\underline{v}}) = \mathbf{cok}_{\boldsymbol{U}}\big(\, \hat{\pi}(\hat{\underline{v}})\,\big).$$

To a C^k-map $F : \boldsymbol{V} \to \boldsymbol{U}$ we can associated a multijet $\mu_k[F]$. This is a C^1-section of the multijet bundle defined by

$$\mu_k[F](\hat{\underline{v}}) = \mathfrak{Proj}_{\hat{\Pi}(\hat{\underline{v}})}\, \boldsymbol{P}_{\hat{\pi}(\hat{\underline{v}})}(F).$$

Note that

$$\dim \mathscr{C}_k(\boldsymbol{V}) = k \dim \boldsymbol{V} = km, \;\; \operatorname{rank} \mathscr{M}_k(\boldsymbol{U}) = km$$

The map F defines a map

$$F^{\times k} : \boldsymbol{V}^k \to \boldsymbol{U}^k, \;\; F^{\times k}(v_1, \dots, v_k) = \big(\, F(v_1, \dots, F(v_k)\,\big) \in \boldsymbol{U}^k.$$

Observe that if $\underline{v} \in \mathscr{C}_k \subset \boldsymbol{V}^k$, then

$$\mathbf{Ev}_{\underline{v}}(F^{\times k}) = \mathbf{Ev}_{\underline{v}}\big(\, \boldsymbol{P}_{\underline{v}} F\,\big).$$

Thus, over $\mathscr{C}_k$, the map $F^{\times k}$ and the map

$$\underline{v} \mapsto \Psi_F(\underline{v}) := \mathbf{Ev}_{\underline{v}}\big(\, \boldsymbol{P}_{\underline{v}} F\,\big) \in \boldsymbol{U}^k$$

have *the same zero sets*. Note also that $\Psi_F(\underline{v}) = 0$ iff $\boldsymbol{P}_{\underline{v}} \in K_{\underline{v}}$, that is, iff

$$\widetilde{F^{\times k}}(\underline{v}) := \mathfrak{Proj}_{\Pi(\underline{v})}\, \boldsymbol{P}_{\underline{v}} F = 0.$$

Thus, over $\mathscr{C}_k$, the maps $F^{\times k}$ and $\widetilde{F^{\times k}}$ have the same zero sets. By definition,

$$\mu_k[F](\hat{\underline{v}}) = \widetilde{F^{\times k}}\big(\, \hat{\pi}(\hat{\underline{v}})\,\big).$$

The map $\hat{\pi}$ restricts to a bijection

$$\{\mu_k(F) = 0\} \cap \widehat{\mathscr{C}}_k^* \to \big\{\, F^{\times k} = 0\,\big\} \cap \mathscr{C}_k.$$

In particular, if $\dim \boldsymbol{V} = \dim \boldsymbol{U}$ and $B \subset \boldsymbol{V}$ is a box, then

$$Z\big[\, B_*^k, F^{\times k}\,\big] = Z\big[\, \hat{\pi}^{-1}(B_*^k), \mu_k(F)\,\big].$$

Let us observe that we have a metric isomorphism of vector bundles

$$\mathcal{I} : \underline{\boldsymbol{U}^k}_{\hat{\pi}^{-1}(B_*^k)} \to \mathcal{M}_k(\boldsymbol{U})\,\big|_{\hat{\pi}^{-1}(B_*^k)},$$

induced by the surjective morphism of product vector bundles,

$$\mathbf{Ev}_{\underline{v}} : \underline{\mathscr{P}^k(\boldsymbol{U})}_{B_*^k} \to \underline{\boldsymbol{U}^k}_{B_*^k}, \;\; \mathcal{I}_{\hat{\underline{v}}} = \big(\, (\mathbf{Ev}_{\underline{v}}\, \mathbf{Ev}_{\underline{v}}^*)^{-1/2}\, \mathbf{Ev}_{\underline{v}}\,\big)^*, \;\; \underline{v} = \hat{\pi}(\hat{\underline{v}}).$$

The bundle isomorphism $\mathcal{I}^{-1} := \mathcal{M}_k(\boldsymbol{U})\,\big|_{B_*^k} \to \underline{\boldsymbol{U}^k}_{B_*^K}$ is the desingularizing renormalization we mention at the end of Subsection 2.4.1.

2.4.5. Higher momentums. Suppose that

$$F : (\Omega, \mathcal{S}, \mathbb{P}) \times \mathscr{V} \to \boldsymbol{U},\ \ (\omega, v) \mapsto F_\omega(v) \in \boldsymbol{U},$$

is a $\boldsymbol{U}$-valued Gaussian random map. We assume that the probability space $(\Omega, \mathcal{S}, \mathbb{P})$ is complete and the map $(\omega, v) \to F_\omega(v)$ is measurable.

The description (2.4.13) of the Kergin projector and the measurability assumption on F show that $(\omega, \underline{v}) \to \boldsymbol{P}_{\underline{v}} F_\omega \in \mathcal{P}^k(\boldsymbol{U})$ is a well defined C^1 *Gaussian* field.

Example 2.4.9. Suppose that $k = 2$. Then

$$\boldsymbol{P}_{\underline{v}}(F)(v) = F(v_0) + F[v_0, v_1](v - v_0) = F(v_0) + \int_0^1 \partial_{(v-v_0)} F\big(\,(1-t)v_0 + tv_1\,\big)dt.$$

If $v_0 \neq v_1$, $r = \|v_1 - v_0\|$, $\nu = \frac{1}{r}(v_1 - v_0)$, then

$$F[v_0, v_1](\nu) = \int_0^1 \partial_\nu F\big(\,v_0 + t(v_1 - v_0)\,\big)dt = \int_0^1 \partial_\nu F\big(\,v_0 + tr\nu\,\big)dt$$

$$= \frac{1}{r}\int_0^r \frac{d}{ds} F\big(\,v_0 + s\nu\,\big)ds = \frac{1}{r}\big(\,F(v_1) - F(v_0)\,\big).$$

We recognize here the vector $\Xi(\underline{v})$ we used in the proof of Proposition 2.4.1. Note that

$$F(v_0) = F(v_1) = 0 \Longleftrightarrow F(v_0) = 0 = F[v_0, v_1]$$

In this case $\mathcal{P}^2(\boldsymbol{U})$ consists of affine maps

$$v \mapsto P(v) = u_0 + Tv,\ \ T \in \operatorname{Hom}(\boldsymbol{V}, \boldsymbol{U}).$$

Then

$$\mathbf{Ev}_{\underline{v}}\, P = \big(\,u_0 + Tv_0,\ \ u_1 + Tv_1\,\big)$$

The

$$P \in K_{\underline{v}} \Longleftrightarrow Tv_0 = Tv_1 = -u_0,\ \ (v_1 - v_0 \in \ker T,\ \ Tv_0 = -u_0$$

The kernel $K_{\underline{v}}(\boldsymbol{U})$ can be identified with $(v_1 - v_0)^\perp \otimes \boldsymbol{U}$, where $(v_1 - v_0)^\perp$ denotes the orthogonal complement in $\boldsymbol{V}$ of the line spanned by $v_1 - v_0$. We have a natural isometric isomorphism $\boldsymbol{U}^k \to K_{\underline{v}}(\boldsymbol{U})^\perp$. □

Lemma 2.4.10. *Let $k \geq 1$. Suppose that $\mathscr{V}$ is an open,* convex *subset of the Euclidean space $\boldsymbol{V}$ and $F : \mathscr{V} \to \boldsymbol{U}$ satisfies the F is* a.s. *C^k and J_{k-1}-ample. Then for any $v \in \mathscr{V}$ there exists an open convex neighborhood $\mathcal{O}_v$ of v in $\mathscr{V}$ such that the restriction of F to $\mathcal{O}_v$ is k-ample.*

Proof. We set

$$[v]_k := (\underbrace{v, \dots, v}_{k}) \in \mathcal{V}^k.$$

Then the Gaussian vector described by the Kergin projector $\boldsymbol{P}_{[v]_k}(F)$ is nondegenerate because it coincides with the degree $k-1$ Taylor polynomial of F at v and this is nondegenerate as a Gaussian vector since F is J_{k-1}-ample.

Since $\boldsymbol{P}_{\underline{v}}$ depends continuously on $\underline{v}$ we deduce that there exists an open convex neighborhood $\mathcal{O}_v$ of v in $\mathscr{V}$ such that, for any $\underline{v} \in \mathcal{O}_v^k$, the Gaussian vector $\boldsymbol{P}_{\underline{v}}(F)$ is nondegenerate. Since the evaluation map

$$\mathbf{Ev}_{\underline{v}} : \mathcal{P}^k(\boldsymbol{U}) \mapsto \boldsymbol{U}^k,\ \mathbf{Ev}_{\underline{v}}(P) = \big(\,P(v_1), \dots, P(v_k)\,\big),$$

is surjective we deduce that the restriction of F to $\mathcal{O}_v$ is k-ample since

$$\mathbf{Ev}_{\underline{v}}\big(\boldsymbol{P}_{\underline{v}}(F)\big) = \big(F(v_1),\dots,F(v_k)\big),\ \ \forall \underline{v}\in\mathscr{V}\setminus\Delta.$$

□

In the remainder of this section I will assume that F is C^k and J_{k-1}-ample.

The *thin* diagonal of $\mathscr{V}^k$, denoted by Δ_0, is the subset

$$\Delta_0 := \big\{\, \underline{v}\in\mathscr{V}^k;\ \ v_1=\cdots=v_k\,\big\}.$$

Equivalently, Δ_0 is the image of $\mathscr{V}$ in $\mathscr{V}^k$ via the diagonal map $u\mapsto[u]_k$. Set

$$\mathcal{O} := \bigcup_{v\in\mathscr{V}}\mathcal{O}_u^k$$

The set $\mathcal{O}$ is an open neighborhood of the thin diagonal and, for any $\underline{v}\in\mathcal{O}$, the Gaussian vector $\boldsymbol{P}_{\underline{v}}(F)$ is nondegenerate.

The multijet random section $\mu_k[F]$ is a.s. C^1. For any $\underline{\tilde{v}}\in\hat{\mathcal{O}}:=\hat{\pi}^{-1}(\mathcal{O})$ the Gaussian vector $\mu_k[F](\underline{\tilde{v}})$ is *nondegenerate* as the image of the nondegenerate vector $\boldsymbol{P}_{\underline{v}}(F)$, $\underline{v}=\hat{\pi}(\underline{\tilde{v}})$, via the linear surjection $\mathscr{P}^k(\boldsymbol{U})\to\mathbf{cok}_{\boldsymbol{U}}(\underline{v})$.

Using the global Kac-Rice formula (2.2.16) we deduce that for any compact set

$$K\subset\hat{\mathcal{O}}:=\hat{\pi}^{-1}(\mathcal{O}),$$

the number of zeros of $\mu_k(F)$ in K has finite mean, i.e.,

$$\mathbb{E}\big[\,Z[K,\mu_k(F)]\,\big]<\infty.$$

Suppose that B is a small box, i.e., a box contained in some $\mathcal{O}_v$. Then $B^k\subset\mathcal{O}$ and the set

$$\widehat{B^k} := \hat{\pi}^{-1}\big(\,B^p\,\big)\subset\hat{\mathcal{O}}$$

is compact. Rercalling the falling factorial notation, $(x)_k = x(x-1)\cdots(x-k+1)$, we deduce

$$\mathbb{E}\big[\,\big(\,Z[B,F]\,\big)_k\,\big] = \mathbb{E}\big[\,Z\big(\,F^{\times p},B_*^k\,\big)\,\big]$$

$$=\mathbb{E}\big[\,Z\big(\,\mu_p[F],\hat{\pi}^{-1}(B_*^p)\,\big)\,\big]\le\mathbb{E}\big[\,Z\big(\,\mu_p[F],\widehat{B^p}\,\big)\,\big]<\infty.$$

In general, for any box $B\subset\mathcal{U}$ there exists a sufficiently fine subdivision $B_i)_{i\in I}$ so that each box of the subdivision is small. Bulinskaya's Lemma implies that

$$Z[B,F]=\sum_{i\in I}Z[B_i,F]\ \text{ a.s.},$$

and we conclude that $Z[B,F]\in L^k$ for any box $B\subset\mathcal{U}$.

We have thus proved the following result.

Theorem 2.4.11. *Let $k\in\mathbb{N}$. Suppose that $\boldsymbol{U},\boldsymbol{V}$ are real Euclidean spaces of the same dimension, $\mathscr{V}\subset\boldsymbol{V}$ is an open set and $F:\mathscr{V}\to\boldsymbol{U}$ is a C^k Gaussian field satisfying the J_{k-1}-ampleness condition*

$$\text{for any } v\in\mathscr{V} \text{ the Gaussian vector } \bigoplus_{j=0}^{k-1}F^{(k)}(v) \text{ is nondegenerate.}\tag{2.4.18}$$

Then for any box $B\subset\mathscr{V}$ we have $Z[B,F]\in L^k$. □

Corollary 2.4.12. *Let $k \in \mathbb{N}$, $k \geq 1$. Suppose that $\boldsymbol{V}$ is real Euclidean spaces of dimension m, $\mathscr{V} \subset \boldsymbol{V}$ is an open set and $\Phi : \mathscr{V} \to \mathbb{R}$ is a C^{k+1} Gaussian field satisfying the J_k-ampleness condition*

$$\text{for any } v \in \mathscr{V} \text{ the Gaussian vector } \bigoplus_{j=0}^{k} F^{(j)}(v) \text{ is nondegenerate.} \tag{2.4.19}$$

Denote by $\mathfrak{C}[B, \Phi]$ the number of critical points of Φ inside the box B. Then $\mathfrak{C}[B, \Phi] \in L^p$. □

Remark 2.4.13. (a) The proof of Theorem 2.4.11 extends to the case of random variables

$$Z[\varphi, F] = \sum_{F(v)=0} \varphi(v), \quad \varphi \in C^0_{\mathrm{cpt}}(\mathscr{V})$$

They are L^k if the assumptions of Theorem 2.4.11 are satisfied.

(b) L. Gass, M. Stecconi [**66**] have given an alternate proof Theorem 2.4.11 that avoids the usage of Hironaka's resolution of singularities theorem, but also relies in a veiled form on the idea of multijet.

(c) The multijet bundle described in this section is a simplified version of the construction of Ancona and Letendre, but it is based on the same technique they introduced in [**4**].

The random multijet $\mu_k[F]$ we described above is nondegenerate only on an open neighborhood $\hat{\mathcal{O}}$ of $\hat{\pi}^{-1}(\Delta_0)$. It is possible that this neighborhood does not contain the entire exceptional locus $\hat{\Delta} = \hat{\pi}^{-1}(\Delta)$.

The more sophisticated multijet constructed in [**4**] is nondegenerate over an open neighborhood of the exceptional locus. This allowed the authors to prove the more refined result, namely, that the expectation of k-th combinatorial momentum of the random measure

$$Z[-, F] = \sum_{F(v)=0} \delta_v$$

(see [**4**, Sec. 6.3]) is a Radon measure over $\mathcal{U}^p$.

The small box localization trick has allowed us to bypass that more sophisticated multijet construction, but we proved an apparently weaker result, namely, for any compactly supported continuous function φ on $\mathcal{U}$ the random variable

$$Z[\varphi, F] = \int_{\mathscr{V}} \varphi(v) \nu_F\big[\, dv \,\big]$$

is k-integrable. However, as shown in [**4**, Prop. 6.25], these properties are equivalent. □

Example 2.4.14. Fix an even Schwartz function $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$ and consider the isotropic Gaussian function $\Phi_{\mathfrak{a}}$ on $\mathbb{R}^m$ introduced in Example 1.2.35. Its spectral measure is

$$\mu_{\mathfrak{a}}\big[\, d\xi \,\big] = \frac{1}{(2\pi)^m} \mathfrak{a}\big(\, |\xi| \,\big)^2 d\xi.$$

As we have seen in Example 1.2.35 this function is a.s. smooth, k-ample and J_k-ample for any $k \in \mathbb{N}$. For any box $B \subset \mathbb{R}^m$ we denote by $\mathfrak{C}_{\mathfrak{a}}[B]$ the number of critical points of $\Phi_{\mathfrak{a}}$ in B. We deduce from Corollary 2.4.12 that $\mathfrak{C}_{\mathfrak{a}}[B] \in L^p$, $\forall p \in [1, \infty)$. □

2.4.6. Some abstract ampleness criteria. We proved that the number of zeros of a Gaussian map has finite k-th momentum assuming two things: the map is C^k and J_{k-1}-ample. The goal of this subsection is to describe some simple guaranteeing various ampleness properties of Gaussian fields. We begin we an abstract technical result that will be be our main tool for detecting ampleness.

Proposition 2.4.15. *Suppose that $\boldsymbol{U}$ is a Banach space with norm $\|-\|$, $\boldsymbol{T}$ is a compact metric space $N \in \mathbb{N}$ and*

$$G : \boldsymbol{U}^N \times \boldsymbol{T} \to [0,\infty), \;\; (u_1,\dots,u_N,t) \mapsto G(u_1,\dots,u_N,t) \in [0,\infty)$$

is a continuous function. We define

$$G_* : \boldsymbol{U}^N \to [0,\infty), \;\; G_*(u_1,\dots,u_N,t) := \min_{t\in\boldsymbol{T}} G(u_1,\dots,u_N,t).$$

Suppose that there exist $v_1,\dots,v_N \in \boldsymbol{U}$ such that $G_*(v_1,\dots,v_N) = r_0 > 0$. Then, for any $r \in (0,r_0)$, there exists $\varepsilon = \varepsilon(r) > 0$ such that

$$\forall u_1,\dots,u_N \in \boldsymbol{U}, \;\; \forall i = 1,\dots N, \;\; \|u_i - v_i\| < \varepsilon \Rightarrow G_*(u_1,\dots,u_N) > r.$$

In particular if

$$U_1 \subset U_2 \subset \cdots$$

is an increasing sequence of finite dimensional subspaces of $\boldsymbol{U}$ whose union is a dense subspace of $\boldsymbol{U}$, then there exists $\nu \in \mathbb{N}$ and

$$u_{1,\nu},\dots,u_{N,\nu} \in U_\nu$$

such that $G_*\big(u_{1,\nu},\dots,u_{N,\nu}\big) > 0$.

Proof. We argue by contradiction. Suppose there exists $r_1 \in (0,r_0)$ and sequences in $\boldsymbol{U}$

$$\big(u_{i,\nu}\big)_{\nu\in\mathbb{N}}, \;\; i = 1,\dots,N,$$

such that

$$\lim_{\nu\to\infty} \|u_{i,\nu} - v_i\| = 0, \;\; \forall i = 1,\dots,N,$$

and

$$G_*(u_{1,\nu},\dots,u_{N,\nu}) \le r_1, \;\; \forall \nu.$$

Next choose $t_\nu \in \boldsymbol{T}$ such that

$$G\big(u_{1,\nu},\dots,u_{N,\nu},t_\nu\big) = G_*\big(u_{1,\nu},\dots,u_{N,\nu}\big)$$

Upon extracting a subsequence we can assume that t_ν converges in $\boldsymbol{T}$ to some point t_∞. Then

$$r_1 \ge \liminf_{\nu\to\infty} G_*\big(u_{1,\nu},\dots,u_{N,\nu}\big) = \liminf_{\nu\to\infty} G\big(u_{1,\nu},\dots,u_{N,\nu},t_\nu\big)$$
$$= G\big(v_1,\dots,v_N,t_\infty\big) \ge r_0 > r_1.$$

□

With $\boldsymbol{T}$ a compact metric space as above, let $E \to \boldsymbol{T}$ be a rank r topological real vector bundle over $\boldsymbol{T}$ equipped with a continuous metric h. For $t \in \boldsymbol{T}$ we denote by $|-|_t$ the norm on the fiber E_t induced by h. The space $C^0(E)$ of continuous sections E is a Banach space with respect to the norm

$$\|u\| := \sup_{t\in\boldsymbol{T}} \big|\, u(t)\,\big|_t, \;\; u \in C(E).$$

Definition 2.4.16. An *ample Banach space of sections* of E is a Banach space $\boldsymbol{U} \subset C^0(E)$ continuously embedded in $C^0(E)$ such that there exist $v_1, \dots, v_N \in \boldsymbol{U}$ such that

$$\forall t \in \boldsymbol{T}, \ \ \operatorname{span}\left\{ u(t), \ \ u \in \boldsymbol{U} \right\} = E_t.$$

Let $k \in \mathbb{N}$. We say that the Banach space $\boldsymbol{U}$ is k-*ample* if for any *distinct* points $t_1, \dots, t_k \in \boldsymbol{T}$ the map

$$\boldsymbol{U} \ni u \mapsto u(t_1) \oplus \cdots \oplus u(t_k) \in E_{t_1} \oplus \cdots \oplus E_{t_k}$$

is onto. □

Example 2.4.17. The space $C^0(E)$ is a k-ample Banach space of continuous sections of $E \to \boldsymbol{T}$ for any $k \in \mathbb{N}$. If $\boldsymbol{T}$ is a compact smooth manifold and $E \to \boldsymbol{T}$ is a smooth vector bundle, then each of the spaces $C^\ell(E)$, $\ell \in \mathbb{N}$, is a k- ample Banach space of sections of E for any $k \in \mathbb{N}$. □

Corollary 2.4.18. *Let $E \to \boldsymbol{T}$ be a real metric vector bundle over the compact metric space $\boldsymbol{T}$ Suppose that $\boldsymbol{U} \subset C^0(E)$ is an ample Banach space of sections*

$$U_1 \subset U_2 \cdots$$

is an increasing sequence of finite dimensional subspaces of $\boldsymbol{U}$ such that

$$U_\infty = \bigcup_{\nu \in \mathbb{N}} U_\nu$$

is dense in $\boldsymbol{U}$. Then there exists $\nu \in \mathbb{N}$, for any $t \in \boldsymbol{T}$, the evaluation map

$$\mathbf{Ev}_t : U_\nu \to E_t \ \text{ is onto.}$$

Proof. Using the compactness of $\boldsymbol{T}$ and the openess of the surjectivuty condition we can find $v_1, \dots, v_N \in \boldsymbol{U}$ such that

$$\forall t \in \boldsymbol{T}, \ \ \operatorname{span}\left\{ v_1(t), \dots, v_N(t) \right\} = E_t.$$

For every $u_1, \dots, u_N \in U$ and $t \in \boldsymbol{T}$ define

$$S_{u_1,\dots,u_N,t} : \mathbb{R}^N \to E_t, \ \ S_{u_1,\dots,u_N,t}(\boldsymbol{x}) = \sum_{k=1}^N x_k u_k(t)$$

and

$$G(u_1, \dots, u_N, t) = \det\left(S_{u_1,\dots,u_N,t} S^*_{u_1,\dots,u_N,t} \right) \geq 0.$$

Note that

$$\operatorname{span}\left\{ u_1(t), \dots, u_N(t) \right\} = E_t \Longleftrightarrow G(u_1, \dots, u_N, t) > 0.$$

Thus

$$G(u_1, \dots, u_N, t) > 0 \iff \mathbf{Ev}_t : \operatorname{span}\left\{ u_1, \dots, u_N \right\} \subset \boldsymbol{U} \to E_t \text{ is onto.}$$

The resulting map $G : \boldsymbol{U}^n \times \boldsymbol{T} \to [0, \infty)$ is continuous and, using the notation in Proposition 2.4.15,

$$G_*(v_1, \dots, v_N) > 0.$$

Using Proposition 2.4.15, we deduce that there exists $\nu \in \mathbb{N}$ and $u_{1,\nu}, \dots, u_{N,\nu} \in U_\nu$ such that

$$G_*(u_{1,\nu}, \dots, u_{N,\nu}) > 0.$$

Hence

$$\mathbf{Ev}_t : \operatorname{span}\{ u_1, \dots, u_N \} \subset \boldsymbol{U} \to E_t \text{ is onto}, \ \ \forall t \in \boldsymbol{T}.$$

A fortiori, this implies that

$$\mathbf{Ev}_t : U_\nu \to E_t \text{ is onto}, \ \ \forall t \in \boldsymbol{T}.$$

□

Corollary 2.4.19. *Let $E \to \boldsymbol{T}$ be a real metric vector bundle over the compact metric space $\boldsymbol{T}$. Suppose that $\boldsymbol{U} \subset C^0(E)$ is a 2-ample Banach space of sections, $U_1 \subset U_2 \cdots$ is an increasing sequence of finite dimensional subspaces of $\boldsymbol{U}$ such that*

$$U_\infty = \bigcup_{\nu \in \mathbb{N}} U_\nu$$

is dense in $\boldsymbol{U}$. Then, for any open neighborhood $\mathcal{O}$ of the diagonal $\Delta \subset \mathbb{T} \times \mathbb{T}$, there exists $\nu \in \mathbb{N}$, for any $(t_1, t_2) \in \boldsymbol{T}^2 \setminus \mathcal{O}$, the evaluation map

$$\mathbf{Ev}_{t_1,t_2} : U_\nu \to E_{t_1} \oplus E_{t_2} \ \ \textit{is onto}.$$

Proof. For $\underline{t} \in \boldsymbol{T}^2$ and $u \in \boldsymbol{U}$ we set

$$u(\underline{t}) := u(t_1) \oplus u(t_2), \ \ E_{\underline{t}} = E_{t_1} \oplus E_{t_2}, \mathbf{Ev}_{\underline{t}}(u) = u(\underline{t}).$$

Using the compactness of $\boldsymbol{T}^2 \setminus \mathcal{O}$ and the openess of the surjectivity condition we deduce that

$$\exists v_1, \dots, v_N \in \boldsymbol{U}, \text{ such that } \ \forall \underline{t} \in \boldsymbol{T}^2 \setminus \mathcal{O}, \ \ \operatorname{span}\{ v_1(\underline{t}), \dots, v_N(\underline{t}) \} = E_{\underline{t}}.$$

For every $u_1, \dots, u_N \in \boldsymbol{U}$ and $\underline{t} \in \boldsymbol{T}^2$ define

$$S_{u_1,\dots,u_N,\underline{t}} : \mathbb{R}^N \to E_{\underline{t}}, \ \ S_{u_1,\dots,u_N,t}(\boldsymbol{x}) = \sum_{k=1}^N x_k u_k(t)$$

and

$$G(u_1, \dots, u_N, \underline{t}) = \det\big(S_{u_1,\dots,u_N,\underline{t}} S^*_{u_1,\dots,u_N,\underline{t}} \big) \geq 0.$$

Note that

$$\operatorname{span}\{ u_1(\underline{t}), \dots, u_N(\underline{t}) \} = E_{\underline{t}} \Longleftrightarrow G(u_1, \dots, u_N, \underline{t}) > 0.$$

Thus

$$G(u_1, \dots, u_N, \underline{t}) > 0 \iff \mathbf{Ev}_t : \operatorname{span}\{ u_1, \dots, u_N \} \subset \boldsymbol{U} \to E_{\underline{t}} \ \text{ is onto}.$$

The resulting map $G : \boldsymbol{U}^n \times \big(\boldsymbol{T}^2 \setminus \mathcal{O} \big) \to [0, \infty)$ is continuous and, using the notation in Proposition 2.4.15, we have $G_*(v_1, \dots, v_N) > 0$.

Proposition 2.4.15, shows that there exists $\nu \in \mathbb{N}$ and $u_{1,\nu}, \dots, u_{N,\nu} \in U_\nu$ such that

$$G_*(u_{1,\nu}, \dots, u_{N,\nu}) > 0.$$

Hence

$$\mathbf{Ev}_{\underline{t}} : \operatorname{span}\{ u_1, \dots, u_N \} \subset \boldsymbol{U} \to E_{\underline{t}} \text{ is onto}, \ \ \forall \underline{t} \in \boldsymbol{T}^2 \setminus \mathcal{O}.$$

A fortiori, this implies that

$$\mathbf{Ev}_{\underline{t}} : U_\nu \to E_{\underline{t}} \text{ is onto}, \ \ \forall \underline{t} \in \boldsymbol{T}^2 \setminus \mathcal{O}.$$

□

Proposition 2.4.20. *Suppose that $E \to \boldsymbol{T}$ is a topological metric vector bundle over the compact metric space $\boldsymbol{T}$. Let $\boldsymbol{X} \subset C^0(E)$ be an ample Banach space of sections of E embedded continuously in $C^0(T)$.*

Suppose that $(u_n)_{n\in\mathbb{N}}$ is a sequence of sections in $\boldsymbol{X}$ such that $\operatorname{span}\{u_n,\ \ n\in\mathbb{N}\}$ *is dense in $\boldsymbol{X}$ and exists $\alpha > 0$ such that*

$$\|u_n\|_{\boldsymbol{U}} = O(n^\alpha) \ \text{ as } n\to\infty. \tag{2.4.20}$$

Fix a sequence of positive real numbers $(\lambda_n)_{n\geq 0}$ such that

$$\liminf_{n\to\infty}\frac{\lambda_n}{n^\beta} > 0, \tag{2.4.21}$$

for some $\beta > 0$. Let $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$ be an even Schwartz function such that $\mathfrak{a}(0) = 1$. Fix a sequence of i.i.d. *standard normal random variables $(X_n)_{n\geq 0}$. Then the following hold.*

(i) *For any $\hbar > 0$ the random series*

$$\sum_{n\in\mathbb{N}}\mathfrak{a}\big(\,\hbar\lambda_n\,\big)X_n u_n \tag{2.4.22}$$

converges a.s. *in $\boldsymbol{X}$. Denote by $\Phi^\hbar$ the resulting continuous Gaussian section of E.*

(ii) *There exists $\hbar_0$ such that $\forall\hbar > \hbar_0$ the Gaussian section $\Phi^\hbar$ is ample.*

Proof. (i) Since $\mathfrak{a}$ is a Schwartz function we deduce from (2.4.20) and (2.4.21) that

$$\sum_{n\to\infty}\big|\,\mathfrak{a}\big(\,\hbar\lambda_n\,\big)\,\big|\|u_n\|_{\boldsymbol{X}} < \infty,\ \ \forall\hbar > 0$$

The convergence of the random series (2.4.22) follows from Proposition 1.1.57.

(ii) For $\hbar > 0$ we set

$$\mathcal{N}_\hbar := \big\{n\in\mathbb{N};\ \ \mathfrak{a}(t\hbar)\neq 0\,\big\}$$

and denote by $\boldsymbol{Y}^\hbar$ the closure in $\boldsymbol{X}$ of

$$\operatorname{span}\big\{\,u_n;\ n\in\mathcal{N}_\hbar\,\big\}.$$

According to Proposition 1.1.57 the above random series defines a *nondegenerate* Gaussian $\boldsymbol{\Gamma}^\hbar$ measure on the Banach space $\boldsymbol{Y}^\hbar$.

Set

$$U_\nu := \operatorname{span}\big\{\,u_1,\dots,u_\nu\,\big\}.$$

Since $\mathfrak{a}(0) = 1$, we deduce that

$$\exists r_0 > 0,\ \ \forall|t|\leq r_0,\ \ \big|\,\mathfrak{a}(t)\,\big|\geq 1/2.$$

Hence, for any $\nu\in\mathbb{N}$ there exists $\hbar = \hbar(\nu) > 0$ so that

$$\forall\hbar\leq\hbar(\nu),\ \ \max_{1\leq k\leq\nu}\hbar\lambda_k < r_0,$$

i.e.,

$$U_\nu\subset\boldsymbol{Y}^\hbar,\ \ \forall\hbar\geq\hbar(\nu).$$

Corollary 2.4.18 implies that there exists $\nu_0\in\mathbb{N}$ such that

$$\forall t\in\boldsymbol{T},\ \ \mathbf{Ev}_t : U_{\nu_0}\to E_t \ \text{ is onto.}$$

Set $\hbar_0 = \hbar(\nu_0)$ such that $U_{\nu_0}\subset\boldsymbol{Y}^\hbar$, $\forall\hbar\leq\hbar_0$.

We will show that for any $t \in \boldsymbol{T}$ and any $\hbar \leq \hbar_0$, the Gaussian vector $\Phi^{\hbar}(t)$ is nondegenerate, i.e., for any open set $\mathcal{O} \subset E_t$, $\mathbb{P}\big[\,\Phi^{\hbar}(t) \in \mathcal{O}\,\big] > 0$. Equivalently, this means

$$\boldsymbol{\Gamma}^{\hbar}\big[\,\mathbf{Ev}_t^{-1}(\mathcal{O})\,\big] > 0.$$

Since $\boldsymbol{\Gamma}^{\hbar}$ is a nondegenerate Gaussian measure on $\boldsymbol{Y}^{\hbar}$, it suffice to show that the open subset $\mathbf{Ev}_t^{-1}(\mathcal{O}) \subset \boldsymbol{Y}^{\hbar}$ is nonempty. This is indeed the case since $\mathbf{Ev}_t^{-1}(\mathcal{O}) \cap U_{\nu_0} \neq \emptyset$. □

Corollary 2.4.21. *Suppose that $E \to M$ is a smooth real vector bundle over the compact smooth manifold M. Fix a smooth Riemann metric g on M, a smooth metric h on E and a smooth connection ∇ on E compatible with h. Let $k \in \mathbb{N}$ and suppose that $(\phi_n)_{n\in\mathbb{N}}$ is a sequence of C^k sections of E that span a dense subset of $C^k(E)$. Suppose that*

$$\|\phi_n\|_{C^k(E)} = O(n^{\alpha}) \;\; \text{as } n \to \infty, \tag{2.4.23}$$

for some $\alpha > 0$. Fix a sequence of positive numbers $(\lambda_n)_{n\in\mathbb{N}}$ satisfying (2.4.21). Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. *standard normal random variables and suppose that $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$ is an even Schwartz function such that $\mathfrak{a}(0) = 1$. Then the following hold.*

(i) *For any $\hbar > 0$ the random series*

$$\sum_{n\in\mathbb{N}} \mathfrak{a}\big(\,\hbar\lambda_n\,\big) X_n \phi_n \tag{2.4.24}$$

converges a.s. *in $C^k(E)$. Denote by $\Phi^{\hbar}$ the resulting C^k Gaussian section of E.*

(ii) *There exists $\hbar_0 > 0$ such that, $\forall \hbar < \hbar_0$, the Gaussian section $\Phi^{\hbar}$ is J_k-ample.*

(iii) *For every point $x \in M$, there exists an open neighborhood $\mathcal{O}_x$ of x in M such that, for any $\hbar < \hbar_0$, the restriction of $\Phi^{\hbar}$ to $\mathcal{O}_x$ is k-ample.*

Proof. (i) This follows from Proposition 2.4.20.

(ii) Consider the jet bundle $J_k(E) \to M$; see (1.2.32). We have a continuous linear

$$C^k(E) \to C^0\big(\,J^k(E)\,\big), \;\; \phi \mapsto J_k(\phi, \nabla).$$

Denote by $\boldsymbol{U}$ the image of this map. It is a closed[8] subspace of $C^0\big(\,J^k(E)\,\big)$. Then the random series

$$\sum_{n\in\mathbb{N}} \mathfrak{a}\big(\,\hbar\lambda_n\,\big) X_n J_k(\phi_n)$$

converges a.s. uniformly to $J_k(\Phi^{\hbar})$. Now observe that $\boldsymbol{U}$ is an ample Banach space of sections of $J^k(E)$. Indeed, using smooth partitions of unity we can find $\psi_1, \dots, \psi_N \in C^k(E)$ such that, for any $x \in M$,

$$\operatorname{span}\big\{\,J_k(\psi_1(x), \dots, J_k(\psi_N)(x)\,\big\} = J_k(E)_x.$$

Proposition 2.4.20 now implies that $J_k(\Phi^{\hbar})$ is an ample Gaussian section of $J_k(E)$. The statement (iii) follows from (ii) by invoking Lemma 2.4.10. □

[8]Here we are using the classical fact that if a sequence of C^1-function (u_n) has the property that both (u_n) and their differentials (du_n) converge uniformly to u and respectively v, then u is C^1 and $du = v$.

Corollary 2.4.22. *Fix an even Schwartz function $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$ and consider the random Fourier series $F^R_{\mathfrak{a}}$ defined in (1.2.21). We regard it as a random smooth function on the torus $\mathbb{T}^m = \big(\mathbb{R}/\mathbb{Z}\big)^m$. Then for any $k \in \mathbb{R}$ there exists $R_k > 0$ such that, for any $R > R_k$ the function $F^R_{\mathfrak{a}}$ is J_k-ample.* □

Lemma 2.4.23. *Suppose that $E \to M$ is a smooth real vector bundle over the compact smooth manifold M. Fix a smooth Riemann metric g on M, a smooth metric h on E and a smooth connection on E compatible with h. Let $k \in \mathbb{N}$ and suppose that $(\phi_n)_{n\in\mathbb{N}}$ is a sequence of C^k sections of E that span a dense subset of $C^k(E)$. Set*

$$U_\nu := \operatorname{span}\big\{\, \phi_1, \dots, \phi_\nu \,\big\}.$$

Then there exists $\nu_0 > 0$ such that $\forall \nu \geq \nu_0$ the following hold.

(i) *For any $t \in M$ and any $\nu \geq \nu_0$ the map*

$$U_\nu \ni u \mapsto J_1(u)_t \in J_1(E_t)$$

is onto. Above, $J_1(u)_t$ is the 1-jet of u at t, $J_1(u)_t = u(t) \oplus \nabla u(t) \in E_t \oplus T^*_t M \otimes E_t$.

(ii) *For any $\underline{t} \in M^2 \setminus \Delta$ the map*

$$U_\nu \ni u \mapsto u(\underline{t}) \in E_{\underline{t}}$$

is onto.

Proof. The space $C^k(E)$ is J_1-ample and arguing as in the proof of Corollary 2.4.18 so there exists $\nu_1 \in \mathbb{N}$ such that for any $\nu \geq \nu_1$ and $t \in M$ the map

$$U_\nu \ni u \mapsto J_1(u)_t \in J_1(E)_t$$

is ample.

The argument at the beginning of Subsection 2.4.5 shows that there exists an open neighborhood $\mathcal{O}$ of the diagonal $\Delta \in M^2$ such that $\forall \nu \geq \nu_1$ and any $\underline{t} \in \mathcal{O} \setminus \Delta$ the map

$$U_\nu \ni u \mapsto u(\underline{t}) \in E_{\underline{t}}$$

is onto.

Corollary 2.4.19 implies that there exists $\nu_0 > 0$ such that $\forall \nu \geq \nu_2$ and any $\underline{t} \in M^2 \setminus \mathcal{O}$ the map

$$U_\nu \ni u \mapsto u(\underline{t}) \in E_{\underline{t}}$$

is onto. Then $\nu_0 = \max(\nu_1, \nu_2)$ has all the claimed properties. □

Corollary 2.4.24. *Fix an even Schwartz function $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$ and consider the random Fourier series $F^\hbar_{\mathfrak{a}}$ defined in (1.2.21). We regard it as a random smooth function on the torus $\mathbb{T}^m = \big(\mathbb{R}/\mathbb{Z}\big)^m$. Then there exists $\hbar = \hbar_{2,2} > 0$ such that, for any $\hbar < \hbar_{2,2}$ the function $F^\hbar_{\mathfrak{a}}$ is J_2-ample and $\nabla F^\hbar_{\mathfrak{a}}$ is 2-ample.* □

2.5. Laws of large numbers

Markov's weak law of large numbers states that if $(X_n)_{n\in\mathbb{N}}$ is a sequence of mean zero, L^2 i.i.d. random variables and

$$S_N = X_1 + \cdots + X_N,$$

then

$$\frac{1}{N} S_N \to 0 \text{ in } L^2.$$

The proof is very simple. The i.i.d. condition shows that

$$\mathrm{Var}\left[\, S_N^2 \,\right] = N\, \mathrm{Var}\left[\, X_1 \,\right].$$

This result can be substantially strengthened by relaxing the i.i.d. assumption to a weak correlation assumption. Namely, the same conclusion is valid if we assume only that there exists a sequence of nonnegative real numbers $(c_k)_{k\geq 0}$ converging to zero such that

$$\mathrm{Cov}\left[\, X_m, X_n \,\right] \leq c\big(\, |m-n| \,\big).$$

In this subsection we prove of a similar result for multiparameter familes of random variables $\big(\, X_{\vec{\ell}}\,\big)_{\vec{\ell}\in\mathbb{N}^m}$.

2.5.1. An abstract law of large numbers for multiparameter familes of random variables. Fix $m\in\mathbb{N}$. Suppose that we have an even continuous function $\rho:\mathbb{R}^m\to(0,\infty)$ that decays sufficiently fast to 0 as $|x|\to\infty$. Then

$$\int_{NB\times NB} \rho(\boldsymbol{x}-\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y} = N^{2m}\int_{B\times B}\rho_N\big(\,\boldsymbol{u}-\boldsymbol{v}\,\big)d\boldsymbol{u}d\boldsymbol{v}$$

where $\rho_N(\boldsymbol{x}) = \rho(N\boldsymbol{x})$. Observing that $\rho_N(\boldsymbol{x})\to 0$ almost everywhere on B we deduce from the dominated convergence theorem that

$$\int_{B\times B}\rho_N\big(\,\boldsymbol{u}-\boldsymbol{v}\,\big)d\boldsymbol{u}d\boldsymbol{v}\to 0$$

as $N\to\infty$. Hence

$$\int_{NB\times NB}\rho(\boldsymbol{x}-\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y} = o\big(\,N^{2m}\,\big) \text{ as } N\to\infty$$

In fact we can be more precise.

If we use Fubini theorem and integrate ρ along the m-planes orthogonal to the 'diagonal $\Delta_N = \{\boldsymbol{x}=\boldsymbol{y}\}\subset NB\times nB$ we deduce that

$$\int_{NB\times NB}\rho(\boldsymbol{x}-\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y} \leq C\,\mathrm{vol}_m\left[\,\Delta_N\,\right]\int_{|\boldsymbol{x}|<N}\rho(\boldsymbol{x})d\boldsymbol{x}\leq CN^m\int_{|\boldsymbol{x}|<N}\rho(\boldsymbol{x})d\boldsymbol{x}.$$

If we specialize further, $\rho(\boldsymbol{x}) = \frac{1}{1+|\boldsymbol{x}|^p}$, $p>0$, $p\neq m$, then

$$\int_{|\boldsymbol{x}|<N}\rho(\boldsymbol{x})d\boldsymbol{x} = O\big(\,N^{\max(m-p,0)}\,\big)$$

so that

$$\int_{NB\times NB}\rho(\boldsymbol{x}-\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y} = O\big(\,N^{m+\max(m-p,0)}\,\big).$$

The sum

$$\sum_{(\vec{k},\vec{\ell})\in\mathbb{I}_N^m\times\mathbb{I}_N^m}\rho(\vec{k}-\vec{\ell})$$

is a very rough Riemann sum approximation of the above integral when $B=[0,1]^N$. The next results show that if $\rho(\boldsymbol{x})=\frac{1}{1+|\boldsymbol{x}|^p}$, then this Riemann sum is also $o\big(N^{2m}\big)$ as $N\to\infty$.

Denote by $|\boldsymbol{x}|_1$ the ℓ^1 norm of $\boldsymbol{x}\in\mathbb{R}^m$,

$$|\boldsymbol{x}|_1:=\sum_{j=1}^m|x_j|.$$

The following technical elementary result is the key to the abstract law of large numbers for multiparameter families.

Lemma 2.5.1. *Fix $m\in\mathbb{N}$. For any $N\in\mathbb{N}$ we set $\mathcal{R}_{N,m}:=\mathbb{I}_N^m\times\mathbb{I}_N^m$.*

(i) *If $m>1$, then there exists a constant $K=K(\alpha,p,m)>0$ such that*

$$\sum_{(\vec{k},\vec{\ell})\in\mathcal{R}_{N_m}}\frac{1}{\big(1+\alpha|\vec{\ell}-\vec{k}|_1\big)^p}\le KN^{2m-\kappa(p)},\;\;\kappa(p)=\min(p,1).$$

(ii) *If $m=1$, then there exists a constant $K=K(\alpha,p)>0$ such that*

$$\sum_{k,\ell\in\mathbb{I}_n}\frac{1}{(1+\alpha|k-\ell|)^p}\le K\begin{cases}N^{2m-\kappa(p)}, & p\neq 1,\\ N\log N, & p=1.\end{cases}$$

Proof. (i) $m>1$. For any $N\in\mathbb{N}$ define

$$D_{N,m}:=\big\{\,(\vec{k},\vec{\ell})\in\mathcal{R}_{N_m};\;\;\exists j=1,\dots,m,\;\;k_j=\ell_j\,\big\},\;\;\mathcal{R}^*_{N,m}:=\mathcal{R}_{N_m}\setminus D_{N,m}.$$

Note that

$$D_N=\bigcup_{i=1}^m D_N^i,\;\;D_N^i=\big\{\,(\vec{k},\vec{\ell})\in\mathcal{R}_N;\;\;k_i=\ell_i\,\big\}.$$

For $1\le i_1<\cdots<i_r\le m$ we have

$$\#\big(\,D_N^{i_1}\cap\cdots\cap D_N^{i_p}\,\big)=N^{2m-2r+1}.$$

Using the Inclusion-Exclusion Principle we deduce that

$$\#D_N=\sum_{r=1}^m(-1)^{p-1}\binom{m}{p}N^{2m-2r+1}\le 2^mN^{2m-1}.$$

We have

$$\sum_{(\vec{k},\vec{\ell})\in\mathcal{R}_N}\frac{1}{\big(1+\alpha|\vec{\ell}-\vec{k}|_1\big)^p}=\underbrace{\sum_{(\vec{k},\vec{\ell})\in D_N}\frac{1}{\big(1+\alpha|\vec{\ell}-\vec{k}|_1\big)^p}}_{=:Y_N}+\underbrace{\sum_{(\vec{k},\vec{\ell})\in\mathcal{R}^*_N}\frac{1}{\big(1+\alpha|\vec{\ell}-\vec{k}|_1\big)^p}}_{=:Z_N}.$$

Note that

$$Y_N\le\#D_N\le 2^mN^{2m-1}.$$

To estimate Z_N we first analyze the structure of the region $\mathcal{R}_N^*$. Denote by R_i the reflection

$$R_i : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m, \quad \begin{bmatrix} x_1 & y_1 \\ \vdots & \vdots \\ \textcircled{x_i} & \boxed{y_i} \\ \vdots & \vdots \\ x_m & y_m \end{bmatrix} \mapsto \begin{bmatrix} x_1 & y_1 \\ \vdots & \vdots \\ \boxed{y_i} & \textcircled{x_i} \\ \vdots & \vdots \\ x_m & y_m \end{bmatrix}.$$

Denote by G_m the direct product of cyclic groups

$$G_m = \big(\mathbb{Z}/2\mathbb{Z}\big)^m = \big\{\, \vec{\epsilon} = (\epsilon_1, \dots, \epsilon_m);\ \ \epsilon_k = 0, 1 \,\big\}.$$

The group G_m acts freely on $\mathcal{R}_N^*$

$$\vec{\epsilon} \cdot \big(\vec{k}, \vec{\ell}\big) = R^{\vec{\epsilon}}\big(\vec{k}, \vec{\ell}\big), \ \ R^{\vec{\epsilon}} = R_1^{\epsilon_1} \cdots R_m^{\epsilon_m}.$$

We denote by $\mathcal{C}_N^+$ the positive chamber of $\mathcal{R}_N^*$,

$$\mathcal{C}_n^+ := \big\{\, \big((\vec{k}, \vec{\ell}\big) \in \mathcal{R}_N^*;\ \ell_j > k_j,\ \ \forall 1 \leq j \leq m \,\big\},$$

and we observe that

$$\mathcal{C}_n^+ = \mathcal{T}_N^m,\ \ \mathcal{T}_N := \big\{\, (k, \ell) \in \mathbb{I}_N \times \mathbb{I}_N;\ \ \ell > k \,\big\}.$$

We have

$$\mathcal{R}_N^* = \bigcup_{\vec{\epsilon} \in G_m} R^{\vec{\epsilon}} \mathcal{C}_n^+.$$

The function ρ is G_m-invariant so

$$\sum_{(\vec{k}, \vec{\ell}) \in \mathcal{R}_N^*} \frac{1}{\big(1 + \alpha|\vec{\ell} - \vec{k}|_1\big)^p} = \sum_{\vec{\epsilon} \in G_m} \sum_{(\vec{k}, \vec{\ell}) \in R^{\vec{\epsilon}} \mathcal{C}_N^+} \frac{1}{\big(1 + \alpha|\vec{\ell} - \vec{k}|_1\big)^p}$$

$$= 2^m \sum_{(\vec{k}, \vec{\ell}) \in \mathcal{C}_N^+} \frac{1}{\big(1 + \alpha|\vec{\ell} - \vec{k}|_1\big)^p} < \frac{2^m}{\alpha^p} \sum_{(\vec{k}, \vec{\ell}) \in \mathcal{C}_N^+} \frac{1}{|\vec{\ell} - \vec{k}|_1^p}$$

(use the AM-GM inequality)

$$\leq \frac{2^m}{(m\alpha)^p} \sum_{(\vec{k}, \vec{\ell}) \in \mathcal{C}_N^+} \Big(\prod_{j=1}^m (\ell_j - k_j)\Big)^{-p/m}.$$

($\mathcal{C}_n^+ = \mathcal{T}_N^m$)

$$= \frac{2^m}{(m\alpha)^p} \prod_{j=1}^m \sum_{(k_j, \ell_j) \in \mathcal{T}_N} \big(\,\ell_j - k_j\,\big)^{-p/m} = \frac{2^m}{(m\alpha)^p} \Big(\sum_{(k, \ell) \in \mathcal{T}_N} \big(\,\ell - k\,\big)^{-p/m}\Big)^m.$$

. Now observe that

$$\sum_{(k, \ell) \in \mathcal{T}_N} \big(\,\ell - k\,\big)^{-p/m} = \sum_{k=1}^{N-1} \sum_{j=1}^{N-k} j^{-p/m}.$$

To proceed further we need to use the following result.

Sublemma 2.5.2. *Let $r \in \mathbb{R}$. Then for any $M \in \mathbb{N}$*

$$S_r(M) := \sum_{j=1}^{M} j^r \leq u_r(M) := \begin{cases} \frac{1}{r+1} M^{r+1}, & r \geq 0, \\ \frac{1}{r+1} M^{r+1} + 1, & r \in (-1, 0), \\ \log M + 1 & r = -1 \\ 2, & r < -1. \end{cases}$$

Proof. Using approximations by Riemann sums for the integral

$$I_r(M) := \int_1^M x^r dx$$

we deduce

$$S_r(M) \leq \begin{cases} I_r(M), & r > 0, \\ I_r(M) + 1, & r < 0 \end{cases} = \begin{cases} \frac{1}{r+1}\big(M^{r+1} - 1\big), & r \geq 0, \\ \frac{1}{r+1}\big(M^{r+1} - 1\big) + 1, & r \in (-1, 0), \\ 1 + \log M, & r = -1, \\ \frac{1}{|r+1|}\big(1 - M^{r+1}\big) + 1, & r < -1. \end{cases}$$

$$\leq \begin{cases} \frac{1}{r+1} M^{r+1}, & r \geq 0, \\ \frac{1}{r+1} M^{r+1}, & r \in (-1, 0), \\ 1 + \log M, & r = -1, \\ 2, & r < -1. \end{cases}$$

□

Suppose that $p \neq m$. Using Sublemma 2.5.2 we deduce that

$$\sum_{j=1}^{N-k} j^{-p/m} \leq u_{-p/m}(N-k) = \begin{cases} \frac{1}{(1-p/m)}(N-k)^{1-p/m} + 1, & p/m < 1, \\ 2 & 1 < p/m. \end{cases}$$

Next, using the sublemma again we deduce

$$\sum_{k=1}^{N} u_r(N-k) \leq \begin{cases} \frac{1}{(1-p/m)(2-p/m)} N^{2-p/m} + N, & p/m < 1, \\ 2N & 1 < p/m. \end{cases}$$

Hence

$$\sum_{(k,\ell) \in \mathfrak{T}_N} \big(\ell - k\big)^{-p/m} \leq \begin{cases} \frac{1}{(1-p/m)(2-p/m)} N^{2-p/m} + N, & p/m < 1, \\ 2N & 1 < p/m. \end{cases}$$

and thus

$$Z_N = \sum_{(\vec{k}, \vec{\ell}) \in \mathcal{R}_N^*} \frac{1}{(1 + |\vec{\ell} - \vec{k}|_1)^p} \leq \frac{2^m}{(m\alpha)^p} \leq C(m, \alpha, p) \begin{cases} N^{2m-p}, & p < m, \\ N^m & p > m. \end{cases} \tag{2.5.1}$$

If $p = m$, then Sublemma 2.5.2 implies that

$$\sum_{j=1}^{N-k} j^{-1} \leq u_{-1}(N-k) = 1 + \log(N-k),$$

and

$$\sum_{k=1}^{N}\big(1+\log(N-k)\big)=N+\log N!.$$

The conclusion follows from Stirling's formula which implies that

$$\log N! = O\big(N\log N\big).$$

(ii) Suppose that $p=1$. Then

$$\sum_{k,\ell=\mathbb{I}_N}\frac{1}{(1+\alpha|k=\ell|)^p}=N+2\sum_{1}\le k<\ell\le n\frac{1}{(1+\alpha|k-\ell|)^p}$$

$$<N+\frac{2}{\alpha^p}\sum_{1\le k<\ell}\frac{1}{(\ell-k)^p}=N+\frac{2}{\alpha^p}\sum_{j=1}^{N-1}\frac{N-j}{j^p}$$

$$=N+\frac{2}{\alpha^p}\sum_{k=1}^{N-1}\sum_{j=1}^{k}\frac{1}{j^p}\le N+\frac{2}{\alpha^p}\sum_{k=1}^{N-1}u_{-p}(j).$$

The conclusion now follows exactly as in (i). □

Lemma 2.5.3. *Fix $m\in\mathbb{N}$. For any $N\in\mathbb{N}$ we set $\mathcal{R}_{N,m}:=\mathbb{I}_N^m\times\mathbb{I}_N^m$. For any $\alpha>0$ and any $p>m$ there exists a constant $K=K(\alpha,p,m)>0$ such that*

$$\sum_{(\vec{k},\vec{\ell})\in\mathcal{R}_{N_m}}\frac{1}{\big(1+\alpha|\vec{\ell}-\vec{k}|_1\big)^p}\le KN^m.$$

Proof. We argue by induction. The case $m=1$ is covered in Lemma 2.5.1 (ii). Define

$$\rho_m:\mathbb{N}^m\times\mathbb{N}^m\to(0,\infty),\ \ \rho_m\big(\vec{k},\vec{\ell}\big)=\frac{1}{\big(1+\alpha|\vec{k}-\vec{\ell}|_1\big)^p}.$$

For any region $\mathcal{R}\subset\mathbb{N}^m\times\mathbb{N}^m$ we set

$$S(\mathcal{R},\rho_m)=\sum_{(\vec{k},\vec{\ell})\in\mathcal{R}}\rho_m\big(\vec{k},\vec{\ell}\big).$$

For any $N\in\mathbb{N}$ define

$$D_{N,m}:=\big\{\,(\vec{k},\vec{\ell})\in\mathcal{R}_{N,m};\ \ \exists j=1,\dots,m,\ \ k_j=\ell_j\,\big\},\ \ \mathcal{R}^*_{N,m}:=\mathcal{R}_{N_m}\setminus D_{N_m}.$$

We have

$$S\big(\mathcal{R}_{N,m},\rho_m\big)=S\big(D_{N,m},\rho_m\big)+S\big(\mathcal{R}^*_{N,m},\rho_m\big).$$

The inequality (2.5.1) implies that

$$S\big(\mathcal{R}^*_{N,m},\rho_m\big)\le KN^m.$$

As before we have

$$D_N=\bigcup_{i=1}^{m}D^i_{N,m},\ \ D^i_{N,m}=\big\{\,(\vec{k},\vec{\ell})\in\mathcal{R}_{N_m};\ \ k_i=\ell_i\,\big\}.$$

Using Inclusion-Exclusion Principle we deduce

$$S\big(D_{N,m},\rho_m\big)=\sum_{p=1}^{m}(-1)^{p-1}\sum_{1\le i_1<\cdots<i_p\le m}S\big(D^{i_1}_{N,m}\cap\cdots\cap D^{i_p}_{N,m},\rho_m\big)$$

$$= \sum_{p=1}^{m} (-1)^{p-1} \binom{m}{p} S\big(D^1_{N,m} \cap \cdots \cap D^p_{N,m}, \rho_m \big)$$

$$= \sum_{p=1}^{m} (-1)^{p-1} \binom{m}{p} S\big(\mathcal{R}_{N,m-p}, \rho_{m-p} \big) \leq \sum_{p=1}^{m} \binom{m}{p} S\big(\mathcal{R}_{N,m-p}, \rho_{m-p} \big)$$

(use the induction assumption)

$$\leq K \sum_{p=1}^{m} \binom{m}{p} N^{m-p} \leq K(N+1)^m \leq 2^m K N^m.$$

□

Corollary 2.5.4. *Consider a family of random variable* $\big(X_{\vec{\ell}} \big)_{\vec{\ell}\in\mathbb{N}^M}$ *defined on the same probability space* $(\Omega, \mathcal{S}, \mathbb{P})$ *such that there exist constants* $C, \alpha, p > 0$, $p \neq m$, *such that*

$$\big|\, \mathrm{Cov}\big[X_{\vec{k}}, X_{\vec{\ell}} \big] \,\big| \leq \frac{C}{\big(1 + \alpha |\vec{k} - \vec{\ell}|_1 \big)^p}, \quad \forall \vec{k}, \vec{\ell} \in \mathbb{N}^m.$$

Then, as $N \to \infty$, *the averages*

$$A_N(X) := \frac{1}{N^m} \sum_{\vec{k}\in\mathbb{I}_N^m} \big(X_{\vec{k}} - \mathbb{E}\big[X_{\vec{k}} \big] \big) \to 0$$

in L^2 *and* a.s..

Proof. Suppose first that $m > m$. Then

$$\mathbb{E}\big[A_N(X)^2 \big] = O\big(N^{-m} \big).$$

If $m > 2$, then

$$\sum_{N\geq 1} N^{-m} < \infty,$$

and we deduce that for any $\varepsilon > 0$

$$\sum_{N\geq 1} \mathbb{P}\big[\,|\, A_N \,|\, > \varepsilon \big] \leq \frac{1}{\varepsilon^2} \sum_{N\geq 1} \big\|\, A_N \,\big\|^2 < \infty,$$

so $A_N \to 0$ a.s.. If $m = 1$ the conclusion follows from [**91**, Thm.10].

If $p < m$ we have $\mathbb{E}\big[A_N(X)^2 \big] = O\big(N^{-1} \big)$. The a.s. convergence follows from the Strong Law of Large Numbers [**104**, Thm. 4]. For $r \in \mathbb{N}$ we denote by C_r the lattice cube $\mathbb{I}^m_{2^r}$. Set $N_r := 2^{r+1}$.

Then

$$u_r := \sum_{\vec{k},\vec{\ell}\in C_{r+1}\setminus C_r} \big|\, \mathbb{E}\big[X_{\vec{k}} X_{\vec{\ell}} \big] \,\big| \leq \sum_{\vec{k},\vec{\ell}\in C_{r+1}} \big|\, \mathbb{E}\big[X_{\vec{k}} X_{\vec{\ell}} \big] \,\big| \leq K N_r^{2m-1},$$

where $K > 0$ is a universal constant. We deduce that

$$\sum_{r\geq 0} \frac{(r+1)^2}{N_r^{2m}} u_r \leq K \sum_{r\geq 0} \frac{(r+1)^2}{N_r} < \infty.$$

According to [**104**, Thm. 4], this implies that $A_N(X) \to 0$ a.s.. □

For latter use we want to mention a version of Corollary 2.5.7 for "pyramidal" arrays.

Corollary 2.5.5. *Let $m \geq 2$. Consider a family of random variable $\big(X^N_{\vec{\ell}}\big)_{N\in\mathbb{N},\,\vec{\ell}\in\mathbb{I}^m_N}$ defined on the same probability space $(\Omega, \mathcal{S}, \mathbb{P})$ such that there exist constants $C, \alpha > 0$, $p > m$ such that*

$$\big|\operatorname{Cov}\big[X^N_{\vec{k}}, X^N_{\vec{\ell}}\big]\big| \leq \frac{C}{\big(1 + \alpha|\vec{k} - \vec{\ell}|_1\big)^p}, \ \ \forall N \in \mathbb{N}\ \ \forall \vec{k}, \vec{\ell} \in bI^m_N.$$

Then, as $N \to \infty$

$$A_N(X) := \frac{1}{N^m}\sum_{\vec{k}\in\mathbb{I}^m_N}\big(X^N_{\vec{k}} - \mathbb{E}\big[X^N_{\vec{k}}\big]\big) \to 0$$

in L^2 and a.s..

Proof. Lemma 2.5.3 implies $\mathbb{E}\big[A_N(X)^2\big] = O(N^{-m})$. Now conclude as in the proof of Corollary 2.5.7. □

2.5.2. SLLN euclid. Fix an even Schwartz function $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$ and consider the isotropic Gaussian function $\Phi_{\mathfrak{a}}$ on $\mathbb{R}^m$ introduced in Example 1.2.35. Its spectral measure is

$$\mu_{\mathfrak{a}}\big[d\xi\big] = \frac{1}{(2\pi)^m}\mathfrak{a}\big(|\xi|\big)^2 d\xi.$$

Its covariance function is determined by

$$\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mathfrak{a}\big(|\xi|\big)^2 d\xi$$

As we have seen in Example 1.2.35 this function is a.s. smooth and k-ample for any $k \in \mathbb{N}$.

For $R > 0$ we set

$$\mathfrak{a}_R(t) := \mathfrak{a}\big(t/R\big), \ \ \forall t \in \mathbb{R}.$$

Consider the finite Borel measure $\mu^R_{\mathfrak{a}} \in \operatorname{Meas}(\mathbb{R}^m)$

$$\mu^R_{\mathfrak{a}}\big[d\xi\big] = \frac{1}{(2\pi)^m} w_{\boldsymbol{a}_R,m}\big(\xi\big)\boldsymbol{\lambda}\big[d\xi\big] = \frac{1}{(2\pi)^m}\mathfrak{a}\big(|\xi|/R\big)^2\boldsymbol{\lambda}\big[d\xi\big].$$

We set $\Phi^R_{\mathfrak{a}} := \Phi_{\mathfrak{a}_R}$. Its covariance kernel is

$$\boldsymbol{K}^R_{\mathfrak{a}}\big(\boldsymbol{x}\big) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mathfrak{a}\big(|\xi|/R\big)^2 d\xi = R^m\boldsymbol{K}_{\mathfrak{a}}\big(R\boldsymbol{x}\big).$$

and we deduce that

$$\Phi^R_{\mathfrak{a}} = R^{m/2}\Phi_{\mathfrak{a}}\big(R\boldsymbol{x}\big).$$

For every $R > 0$ the random function $\Phi^R_{\mathfrak{a}}$ is a.s. Morse and there is an associated critical random measure

$$\mathfrak{C}^R_{\mathfrak{a}} := \sum_{\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})=0}\delta_{\boldsymbol{x}}.$$

Thus, for any Borel subset $S \subset \mathbb{R}^m$, $\mathfrak{C}^R_{\mathfrak{a}}[S]$ is the number of critical points of $\Phi^R_{\mathfrak{a}}$ in S. More generally, if $f : \mathbb{R}^m \to \mathbb{R}$ is a continuous compactly supported function we set

$$\mathfrak{C}^R_{\mathfrak{a}}[f] = \int_{\mathbb{R}^m} f(\boldsymbol{x})\mathfrak{C}^R_{\mathfrak{a}}\big[d\boldsymbol{x}\big] = \sum_{\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})=0} f(\boldsymbol{x}).$$

Let $\mathfrak{C}_{\mathfrak{a}} := \mathfrak{C}^R_{\mathfrak{a}}\big|_{R=1}$. The main goal of this subsection is to investigate the behavior of $\mathfrak{C}^R_{\mathfrak{a}}$ in the white noise limit, $R \nearrow \infty$. The main result is the following.

Theorem 2.5.6. *Fix an amplitude $\mathfrak{a} \in \mathcal{S}(\mathbb{R})$. Then the following hold.*

(i) *There exists a universal explicit constant $C_m(\mathfrak{a}) > 0$ such that for any $f \in C^0_{\rm cpt}(\mathbb{R}^m)$ we have*

$$\mathbb{E}\big[\,\mathfrak{C}^R_{\mathfrak{a}}[f]\,\big] = C_m(\mathfrak{a}) R^m \int_{\mathbb{R}^m} f(\boldsymbol{x}) d\boldsymbol{x}.$$

(ii) *There exists a constant $V_m(\mathfrak{a}) \geq 0$, that depends only on m and $\mathfrak{a}$ such that for any $f \in C^0_{\rm cpt}(\mathbb{R}^m)$*

$$\operatorname{Var}\big[\,\mathfrak{C}^R_{\mathfrak{a}}[f]\,\big] \sim V_m(\mathfrak{a}) R^m \int_{\mathbb{R}^m} f(\boldsymbol{x})^2 d\boldsymbol{x}, \;\; \textit{as } R \to \infty.$$

□

The case $m = 1$ of this theorem was proved by M. Ancona and T. Letendre [**3**, Thm. 1.16] and L. Gass [**65**, Thm.1.6]. This theorem was recently proved by a different method in [**2**]. One immediate application of Theorem 2.5.6 is a law of large numbers.

For any positive integer N we denote by $\mathcal{L}_N$ the random measure

$$\mathcal{L}_N := \frac{1}{N^m} \mathfrak{C}^N_{\mathfrak{a}}.$$

Theorem 2.5.6 shows for any $f \in C^0_{\rm cpt}(\mathbb{R}^m)$ we have

$$\operatorname{Var}\big[\,\mathcal{L}_N[f]\,\big] \sim \mathit{const} \times N^{-m} \;\; \text{as } N \to \infty$$

Since $\sum_{n \geq 1} n^{-m} < \infty$ for $m \geq 2$, we deduce from Borel-Cantelli the following law of large numbers.

Corollary 2.5.7. *Let $m \geq 2$. Then for any $f \in C^0_{\rm cpt}(\mathbb{R}^m)$,*

$$\lim_{N\to\infty} \mathcal{L}_N[f] = C_m(\mathfrak{a}) \boldsymbol{\lambda}\big[\, f\,\big] = C_m(\mathfrak{a}) \int_{\mathbb{R}^m} f(\boldsymbol{x}), \;\; \text{a.s. } \textit{and } L^2, \tag{2.5.2}$$

where $\boldsymbol{\lambda}$ denotes the Lebesgue measure. □

Proof of Theorem 2.5.6. . We need to introduce some notation. Set

- Define
$$\widehat{\Phi} : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}, \;\; \widehat{\Phi}(\boldsymbol{x}, \boldsymbol{y}) = \Phi_{\mathfrak{a}}(\boldsymbol{x}) + \Phi_{\mathfrak{a}}(\boldsymbol{y}), \;\; \hat{\mathfrak{C}} = \mathfrak{C}_{\widehat{\Phi}},$$
$$\widehat{H}(\boldsymbol{x}, \boldsymbol{y}) := \operatorname{Hess}_{\widehat{\Phi}}(\boldsymbol{x}, \boldsymbol{y}), \;\; H(\boldsymbol{x}) := \operatorname{Hess}_{\Phi_{\mathfrak{a}}}(\boldsymbol{x}).$$
- Choose an independent copy $\Psi_{\mathfrak{a}}$ of $\Phi_{\mathfrak{a}}$ and set
$$\widetilde{\Phi}(\boldsymbol{x}, \boldsymbol{y}) := \Phi_{\mathfrak{a}}(\boldsymbol{x}) + \Psi_{\mathfrak{a}}(\boldsymbol{y}), \;\; \widetilde{H}(\boldsymbol{x}, \boldsymbol{y}) := \operatorname{Hess}_{\widetilde{\Phi}}(\boldsymbol{x}, \boldsymbol{y}), \;\; \tilde{\mathfrak{C}} = \mathfrak{C}_{\widetilde{\Phi}}.$$
- Set $\|f\| := \|f\|_{C^0(\mathbb{R}^m)}$.
- Set
$$\mathfrak{X} = \mathbb{R}^m \times \mathbb{R}^m \setminus \Delta = \big\{\, (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^m \times \mathbb{R}^m; \;\; \boldsymbol{x} \neq \boldsymbol{y} \,\big\}.$$

Observe that the random function on $\widehat{\Phi}(\boldsymbol{x}, \boldsymbol{y})$ is *stationary* with respect to the action of $\mathbb{R}^m$ on $\mathbb{R}^m \times \mathbb{R}^m$ itself by translations

$$T_{\boldsymbol{v}}(\boldsymbol{x}, \boldsymbol{y}) = \big(\, \boldsymbol{x} + \boldsymbol{v}, \boldsymbol{y} + \boldsymbol{v}\,\big), \;\; \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{v} \in \mathbb{R}^m, \tag{2.5.3}$$

whereas $\widetilde{\Phi}$ is stationary with respect to the action by translations of $\mathbb{R}^{2m}$ on itself.

We have

$$\hat{\mathfrak{C}}[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}] = \sum_{\substack{\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})=\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})=0,\\ \boldsymbol{x}\neq\boldsymbol{y}}} f_R(\boldsymbol{x})f_R(\boldsymbol{y}) = \mathfrak{C}_{\mathfrak{a}}[f_R]^2 - \mathfrak{C}_{\mathfrak{a}}[f_R^2].$$

Bulinskaya's lemma implies that

$$\mathbb{P}\big[\,\exists\boldsymbol{x}:\;\; \nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}) = \nabla\Psi_{\mathfrak{a}}(\boldsymbol{x}) = 0\,\big] = 0$$

and we deduce

$$\tilde{\mathfrak{C}}[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}] = \sum_{\substack{\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})=\nabla\Psi_{\mathfrak{a}}(\boldsymbol{y})=0,\\ \boldsymbol{x}\neq\boldsymbol{y}}} f_R(\boldsymbol{x})f_R(\boldsymbol{y})$$

$$= \sum_{\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})=\nabla\Psi_{\mathfrak{a}}(\boldsymbol{y})=0} f_R(\boldsymbol{x})f_R(\boldsymbol{y}) = \mathfrak{C}[f,\Phi_{\mathfrak{a}}]\,]\mathfrak{C}[f,\Psi_{\mathfrak{a}}], \;\; \text{a.s..}$$

Hence

$$\mathbb{E}\big[\,\mathfrak{C}[f_R,\Phi_{\mathfrak{a}}]\cdot\mathfrak{C}[f,\Psi_{\mathfrak{a}}]\,\big] = \mathbb{E}\big[\,\mathfrak{C}[f_R,\Phi_{\mathfrak{a}}]\,\big]\;\cdot\;\mathbb{E}\big[\,\mathfrak{C}[f_R,\Psi_{\mathfrak{a}}]\,\big] = \mathbb{E}\big[\,\mathfrak{C}[f_R,\Phi_{\mathfrak{a}}]\,\big]^2$$

so that

$$\mathbb{E}\big[\,\hat{\mathfrak{C}}^R[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}]\,\big] - \mathbb{E}\big[\,\tilde{\mathfrak{C}}^R[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}]\,\big] = \underbrace{\mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[f_R]^2\,\big] - \mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[f_R]\,\big]^2}_{=\mathrm{Var}\,\big[\,\big[\,\mathfrak{C}_{\mathfrak{a}}[f_R]\,\big]} - \mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[f_R^2]\,\big]. \tag{2.5.4}$$

We have seen that

$$\mathbb{E}\big[\,\mathfrak{C}_{\mathfrak{a}}[f_R^2]\,\big] = R^m C_m(\mathfrak{a})\int_{\mathbb{R}^m} f^2(\boldsymbol{x})d\boldsymbol{x}$$

so we have to show that

$$I(R) := \mathbb{E}\big[\,\hat{\mathfrak{C}}[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}]\,\big] - \mathbb{E}\big[\,\tilde{\mathfrak{C}}[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}]\,\big] \sim Z_m(\mathfrak{a})R^m\int_{\mathbb{R}^m} f(\boldsymbol{x})^2 d\boldsymbol{x} \;\text{ as } R\to\infty \tag{2.5.5}$$

for some constant $Z_m(\mathfrak{a})\in\mathbb{R}$ that depends only on m and $\mathfrak{a}$.

Lemma 2.5.8. *For any $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^m$, $\boldsymbol{x}\neq\boldsymbol{y}$, the Gaussian vector $\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})$ is nondegenerate.*

Proof. We have

$$\mathrm{Var}\,\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix} \mathrm{Var}\,\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \mathrm{Cov}\,\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big] \\ & \\ \mathrm{Cov}\,\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \mathrm{Var}\,\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big] \end{bmatrix}.$$

As shown in (2.3.20), for any $\boldsymbol{x}\in\mathbb{R}^n$ we have

$$\mathrm{Var}\,\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] = d_m\mathbb{1}_m,\;\; d_m = \int_{\mathbb{R}^n}\xi_1^2\mu_{\mathfrak{a}}\big[\,d\xi\,\big].$$

We have

$$\mathrm{Cov}\,\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big] = \big(\,\partial_{x_j}\partial_{y_k}\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}-\boldsymbol{y})\,\big)_{1\le j,k\le m}$$

and

$$\partial_{x_j}\partial_{y_k}\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}-\boldsymbol{y}) = \int_{\mathbb{R}^m} e^{-\boldsymbol{i}\langle\xi,\boldsymbol{x}-\boldsymbol{y}\rangle}\xi_j\xi_k\mu_{\mathfrak{a}}\big[\,d\xi\,\big]. \tag{2.5.6}$$

Since $\Phi_{\mathfrak{a}}$ is stationary it suffice to consider only the case $\boldsymbol{x} = 0$. On the other hand, $\Phi_{\mathfrak{a}}$ is $O(m)$-invariant so, up to a rotation, we can assume that $\boldsymbol{x} - \boldsymbol{y} = -t\boldsymbol{e}_1$, $t \neq 0$, where $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ is the canonical basis of $\mathbb{R}^m$. Hence

$$\partial_{x_j}\partial_{y_k}\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}-\boldsymbol{y}) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}t\xi_1}\xi_j\xi_k\mu_{\mathfrak{a}}\big[\,d\xi\,\big].$$

Let us observe that if $j \neq k$, then either $j \neq 1$, or $k \neq 1$. Suppose $j \neq 1$. The function $e^{\boldsymbol{i}t\xi_1}\xi_j\xi_k$ is odd with respect to the reflection $\xi_j \mapsto -\xi_j$ so

$$\partial_{x_j}\partial_{y_k}\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x},\boldsymbol{y}) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}t\xi_1}\xi_j\xi_k\mu_{\mathfrak{a}}\big[\,d\xi\,\big] = 0,\ \ \forall j \neq k.$$

If $j = k$, then

$$d_m(j) := \partial_{x_j}\partial_{y_j}\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x},\boldsymbol{y}) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}t\xi_1}\xi_j^2\mu_{\mathfrak{a}}\big[\,d\xi\,\big] = \int_{\mathbb{R}^m} \cos(t\xi_1)\xi_j^2\mu_{\mathfrak{a}}\big[\,d\xi\,\big]$$

and we deduce[9]

$$|v_m(j)| \leq \int_{\mathbb{R}^m} \big|\cos(t\xi_1)\big|\xi_j^2\mu_{\mathfrak{a}}\big[\,d\xi\,\big] < \int_{\mathbb{R}^m} \xi_j^2\mu_{\mathfrak{a}}\big[\,d\xi\,\big] = d_m.$$

After a reordering

$$\begin{gathered}\big(\,\partial_{x_1}\Phi_{\mathfrak{a}}(\boldsymbol{x}),\dots,\partial_{x_m}\Phi_{\mathfrak{a}}(\boldsymbol{x}),\partial_{y_1}\Phi_{\mathfrak{a}}(\boldsymbol{y}),\dots\partial_{y_m}\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big)\\ \downarrow\\ \big(\,\partial_{x_1}\Phi_{\mathfrak{a}}(\boldsymbol{x}),\partial_{y_1}\Phi_{\mathfrak{a}}(\boldsymbol{y}),\dots,\partial_{x_m}\Phi_{\mathfrak{a}}(\boldsymbol{x}),\partial_{y_m}\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big)\end{gathered}$$

we see that

$$\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big] = \bigoplus_{j=1}^m \underbrace{\begin{bmatrix} d_m & d_m(j)\\ d_m(j) & d_m\end{bmatrix}}_{=:V_j}.$$

Note that, for each j, the symmetric matrix V_j is positive definite since

$$\det V_j = d_m^2 - d_m(j)^2 > 0.$$

□

Since $\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})$ is nondegenerate for $\boldsymbol{x} \neq \boldsymbol{y}$ we can use the Kac-Rice formula to compute $\mathbb{E}\big[\,\hat{\mathfrak{C}}^R[\boldsymbol{I}_{\mathfrak{X}}f_R^{\boxtimes 2}]\,\big]$. We deduce that for any $R > 0$

$$\begin{gathered}\mathbb{E}\big[\,\hat{\mathfrak{C}}[\boldsymbol{I}_{\mathfrak{X}}f_R^{\boxtimes 2}]\,\big]\\ = \int_{\mathbb{R}^m\times\mathbb{R}^m\setminus\Delta} \underbrace{\mathbb{E}\big[\,|\det\widehat{H}(\boldsymbol{x},\boldsymbol{y})|\,\big|\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y}) = 0\,\big]p_{\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})}(0)}_{=\widehat{\rho}(\boldsymbol{x},\boldsymbol{y})} f_R^{\boxtimes 2}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}d\boldsymbol{y}\,\big].\end{gathered} \tag{2.5.7}$$

Since $\widehat{\Phi}$ is invariant under the translations (2.5.3) we deduce that $\widehat{\rho}$ depends only on $\boldsymbol{x}-\boldsymbol{y}$.

[9] At this point we use the fact that $\mathfrak{a}\big(\,|\xi|\,\big) > 0$ for $|\xi|$ sufficiently small.

The gradient $\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})$ is nondegenerate for any $\boldsymbol{x},\boldsymbol{y}$ and invoking Kac-Rice again we obtain

$$\begin{aligned}
&\mathbb{E}\big[\,\tilde{\mathfrak{C}}[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}]\,\big]\\
&=\int_{\mathbb{R}^m\times\mathbb{R}^m\setminus\Delta}\underbrace{\mathbb{E}\big[\,|\det\widetilde{H}(\boldsymbol{x},\boldsymbol{y})|\,\big|\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})=0\,\big]p_{\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})}(0)}_{=\widetilde{\rho}(\boldsymbol{x},\boldsymbol{y})}f_R^{\boxtimes 2}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}d\boldsymbol{y}\,\big].
\end{aligned}\tag{2.5.8}$$

The function $\widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})$ is independent of $\boldsymbol{x},\boldsymbol{y}$ since the random function $\widetilde{\Phi}$ is stationary. Thus

$$I(R)=\int_{\mathfrak{X}}\big(\underbrace{\widehat{\rho}(\boldsymbol{x},\boldsymbol{y})-\widetilde{\rho}(\boldsymbol{x},\boldsymbol{y})}_{=:\Delta(\boldsymbol{x},\boldsymbol{y})}\big)f_R(\boldsymbol{x})f_R(\boldsymbol{y})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}d\boldsymbol{y}\,\big].\tag{2.5.9}$$

There is a serious issue concerning $\widehat{\rho}(\boldsymbol{x},\boldsymbol{y})$ namely it blows up as $(\boldsymbol{x},\boldsymbol{y})$ approaches the diagonal so this Kac-Rice density may not be locally integrable.

We first describe the behavior of $\Delta(\boldsymbol{x},\boldsymbol{y})$ away from the diagonal. Note that Δ depends only on the $\boldsymbol{x}-\boldsymbol{y}$.

For every $\boldsymbol{z}\in\mathbb{R}^m$ we set

$$T(z):=\sum_{|\alpha|\leq 4}\big|\,\partial^\alpha\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{z})\,\big|.$$

Since $\boldsymbol{K}_{\mathfrak{a}}$ is a Schwartz function we deduce that

$$T(\boldsymbol{z})=O\big(\,|\boldsymbol{z}|^{-\infty}\,\big)\ \text{ as } |\boldsymbol{z}|\to\infty.$$

This means that

$$\forall p>0,\ \ T(\boldsymbol{z})=O\big(\,|\boldsymbol{z}|^{-p}\,\big)\ \text{ as } |\boldsymbol{z}|\to\infty.$$

Observe that

$$\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]=\begin{bmatrix}\operatorname{Var}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & 0\\ 0 & \operatorname{Var}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]\end{bmatrix}=d_m\mathbb{1}_{2m},$$

and

$$\begin{aligned}
\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]&=\begin{bmatrix}\operatorname{Var}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]\\ \\ \operatorname{Cov}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Var}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]\end{bmatrix}\\
&=\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]+\underbrace{\begin{bmatrix}0 & \operatorname{Cov}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]\\ \\ \operatorname{Cov}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & 0\end{bmatrix}}_{=:R_\nabla(\boldsymbol{x},\boldsymbol{y})}.
\end{aligned}$$

Hence

$$\big\|\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]-\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]\big\|_{\mathrm{op}}=\|R_\nabla(\boldsymbol{x},\boldsymbol{y})\|_{\mathrm{op}}=O\big(\,T_R(\boldsymbol{x}-\boldsymbol{y})\,\big),\tag{2.5.10}$$

The operators $\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]$ and $\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]$ are invertible for $\boldsymbol{x}\neq\boldsymbol{y}$. In particular

$$\begin{aligned}
\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}&=\Big(\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]+R_\nabla(\boldsymbol{x},\boldsymbol{y})\Big)^{-1}\\
&=\Big(\mathbb{1}+\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}R_\nabla(\boldsymbol{x},\boldsymbol{y})\Big)^{-1}\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1},
\end{aligned}\tag{2.5.11}$$

$$\big\|\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}-\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\big\|_{\mathrm{op}}=O\big(\,T(\boldsymbol{x}-\boldsymbol{y})\,\big)\text{ as }|\boldsymbol{x}-\boldsymbol{y}|\to\infty.\tag{2.5.12}$$

Note that

$$\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})\,\big]=\left[\begin{array}{cc}\operatorname{Var}\big[\,H(\boldsymbol{x})\,\big] & 0\\ 0 & \operatorname{Var}\big[\,H(\boldsymbol{y})\,\big]\end{array}\right].$$

Since $\Phi_{\mathfrak{a}}$ is stationary, $\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})\,\big]$ is *independent* of $\boldsymbol{x}$ and $\boldsymbol{y}$. We have

$$\operatorname{Var}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y})\,\big]=\left[\begin{array}{cc}\operatorname{Var}\big[\,H(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,H(\boldsymbol{x}),H(\boldsymbol{y})\,\big]\\ & \\ \operatorname{Cov}\big[\,H(\boldsymbol{y}),H(\boldsymbol{x})\,\big] & \operatorname{Var}\big[\,H(\boldsymbol{y})\,\big]\end{array}\right]$$

$$=\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})\,\big]+\underbrace{\left[\begin{array}{cc}0 & \operatorname{Cov}\big[\,H(\boldsymbol{x}),H(\boldsymbol{y})\,\big]\\ & \\ \operatorname{Cov}\big[\,H(\boldsymbol{y}),H(\boldsymbol{x})\,\big] & 0\end{array}\right]}_{=:R_H(\boldsymbol{x},\boldsymbol{y})}.$$

We deduce that as $|\boldsymbol{x}-\boldsymbol{y}|\to\infty$ we have

$$\Big\|\,\operatorname{Var}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y})\,\big]-\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})\,\big]\,\Big\|_{\mathrm{op}}=\|R_H(\boldsymbol{x},\boldsymbol{y})\|_{\mathrm{op}}=O\big(\,T(\boldsymbol{x}-\boldsymbol{y})\,\big).\tag{2.5.13}$$

We denote by $\widetilde{H}(\boldsymbol{x},y)^{\flat}$ the Gaussian random matrix

$$\widetilde{H}(\boldsymbol{x},\boldsymbol{y})^{\flat}=\widetilde{H}(\boldsymbol{x},\boldsymbol{y})-\mathbb{E}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})\,\|\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y}\,\big].$$

Similarly, we denote by $\widehat{H}(\boldsymbol{x},\boldsymbol{y})^{\flat}$ the Gaussian random matrix

$$\widehat{H}(\boldsymbol{x},y)^{\flat}=\widehat{H}(\boldsymbol{x},\boldsymbol{y})-\mathbb{E}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y})\,\|\,,\nabla\widehat{\Phi}\,\big].$$

The distributions of $\widetilde{H}(\boldsymbol{x},\boldsymbol{y})^{\flat}$ and $\widehat{H}(\boldsymbol{x},\boldsymbol{y})^{\flat}$ are determined by the Gaussian regression formula (1.1.18).

Since $\widetilde{H}(\boldsymbol{x},y)$ and $\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})$ are independent we deduce

$$\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})^{\flat}\,\big]=\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})\,\big].$$

Using the regression formula we deduce that for $|\boldsymbol{x}-\boldsymbol{y}|>C_0$,

$$\begin{gathered}\operatorname{Var}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y})^{\flat}\,\big]=\operatorname{Var}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y})\,\big]\\ -\operatorname{Cov}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\operatorname{Cov}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y}),\widehat{H}(\boldsymbol{x},\boldsymbol{y})\,\big]|\\ =\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})^{\flat}\,\big]+R_H(\boldsymbol{x},\boldsymbol{y})\\ -\operatorname{Cov}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\operatorname{Cov}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y}),\widehat{H}(\boldsymbol{x},\boldsymbol{y})\,\big].\end{gathered}$$

We have

$$\operatorname{Cov}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]=\left[\begin{array}{cc}\operatorname{Cov}\big[\,H(\boldsymbol{x}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,H(\boldsymbol{x}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]\\ & \\ \operatorname{Cov}\big[\,H(\boldsymbol{y}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,H(\boldsymbol{y}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]\end{array}\right]$$

$$=\left[\begin{array}{cc}\operatorname{Cov}\big[\,0 & \operatorname{Cov}\big[\,H(\boldsymbol{x}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{y})\,\big]\\ & \\ \operatorname{Cov}\big[\,H(\boldsymbol{y}),\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] & 0\end{array}\right].$$

This implies

$$\operatorname{Cov}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]=O\big(\,T(\boldsymbol{x}-\boldsymbol{y})\,\big)\ \text{ as } |\boldsymbol{x}-\boldsymbol{y}|\to\infty.$$

We deduce from (2.5.12) that

$$\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}=\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}+O\big(\,T(\boldsymbol{x}-\boldsymbol{y})\,\big).$$

Since $\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]$ is independent of $\boldsymbol{x}$ and $\boldsymbol{y}$ we conclude that

$$\operatorname{Cov}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\operatorname{Cov}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y}),\widehat{H}(\boldsymbol{x},\boldsymbol{y})\,\big]=O\big(\,T(\boldsymbol{x}-\boldsymbol{y})\,\big),\tag{2.5.14}$$

Since $\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})\,\big]$ is independent of $\boldsymbol{x},\boldsymbol{y}$ we deduce that there exists $\mu_0>0$ such that

$$\operatorname{Var}\big[\,\widetilde{H}(\boldsymbol{x},\boldsymbol{y})^\flat\,\big]\geq\mu_0\mathbb{1},\ \ \forall\boldsymbol{x}\neq\boldsymbol{y}.$$

We deduce from (2.5.14) and Lemma 1.1.27 that

$$\Big|\,\mathbb{E}\big[\,|\det\widehat{H}(\boldsymbol{x},\boldsymbol{y})^\flat|\,\big]-\mathbb{E}\big[\,|\det\widetilde{H}(\boldsymbol{x},\boldsymbol{y})^\flat|\,\big]\,\Big|=O\big(\,T(\boldsymbol{x}-\boldsymbol{y})^{1/2}\,\big).\tag{2.5.15}$$

Using (2.5.12) we deduce that as $|\boldsymbol{x}-\boldsymbol{y}|\to\infty$ we have

$$\begin{gathered}\Big|\,p_{\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})}(0)-p_{\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})}(0)\,\Big|\\=(2\pi)^{-m/2}\Big|\,\det\operatorname{Var}\big[\,\nabla\widehat{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}-\det\operatorname{Var}\big[\,\nabla\widetilde{\Phi}(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\,\Big|=O\big(\,T(\boldsymbol{x}-\boldsymbol{y})\,\big).\end{gathered}\tag{2.5.16}$$

Note also that (2.5.13) implies that

$$\sup_{|\boldsymbol{x}-\boldsymbol{y}|_\infty>1}\|\operatorname{Var}\big[\,\widehat{H}(\boldsymbol{x},\boldsymbol{y})^\flat\,\big]\|_{\mathrm{op}}<\infty\tag{2.5.17}$$

We can now estimate the right-hand-side of (2.5.33). Using (2.5.15), (2.5.16) and (2.5.17) we conclude that

$$\forall|\boldsymbol{x}-\boldsymbol{y}|>1,\ \ \big|\,\Delta(\boldsymbol{x},\boldsymbol{y})\,\big|=O\big(\,\big|\,\boldsymbol{x}-\boldsymbol{y}\,\big|^{-\infty}\,\big),\ \ \text{as } |\boldsymbol{x}-\boldsymbol{y}|\to\infty.\tag{2.5.18a}$$

$$\sup_{|\boldsymbol{x}-\boldsymbol{y}|>1}\big|\,\Delta(\boldsymbol{x}-\boldsymbol{y})\,\big|<\infty.\tag{2.5.18b}$$

Suppose that

$$\operatorname{supp}f\subset\{|\boldsymbol{x}|\leq_0\}.$$

Denote by $\widehat{\mathfrak{X}}$ the radial-blowup of $\mathbb{R}^m\times\mathbb{R}^m$ along the diagonal. It is diffeomorphic to the product $\mathbb{R}^m\times S^{m-1}\times[0,\infty)$.

Choose new orthogonal coordinates (ξ,η) given by

$$\xi=\boldsymbol{x}+\boldsymbol{y},\ \ \eta=\boldsymbol{x}-\boldsymbol{y}\Longleftrightarrow\boldsymbol{x}=\frac{1}{2}(\xi+\eta),\ \ \boldsymbol{y}=\frac{1}{2}(\xi-\eta)$$

then

$$|\boldsymbol{x}-\boldsymbol{y}|=|\eta|,\ \ d\boldsymbol{x}d\boldsymbol{y}=2^{-2m}d\xi d\eta.$$

Recall that $\widehat{\rho}$ depends only on η. Note that if $\boldsymbol{x},\boldsymbol{y}\in\operatorname{supp}f$, then $|\boldsymbol{x}|,|\boldsymbol{y}|<r_0$ and thus

$$\boldsymbol{x},\boldsymbol{y}\in\operatorname{supp}f\ \Rightarrow\ |\xi|,\ |\eta|<\frac{1}{2}|\xi+\eta|+\frac{1}{2}|\xi-\eta|=|\boldsymbol{x}|+|\boldsymbol{y}|\leq 2r_0.\tag{2.5.19}$$

The natural projection $\pi:\widehat{\mathfrak{X}}\to\mathbb{R}^m\times\mathbb{R}^m$ can given the explicit description

$$\mathbb{R}^m\times S^{m-1}\times[0,\infty)\ni(\xi,\boldsymbol{\nu},r)\mapsto(\xi,\eta)=(\xi,r\boldsymbol{\nu})\in\mathbb{R}^m\times\mathbb{R}^m.$$

We set

$$w(\boldsymbol{x},\boldsymbol{y})=|\boldsymbol{x}-\boldsymbol{y}|^{m-2}\widehat{\rho}(\boldsymbol{x},\boldsymbol{y}).$$

We deduce from [16, Sec. 4.2] or [55, Appendix A.1]

$$\sup_{0<|\boldsymbol{x}-\boldsymbol{y}|\leq 1} \bigl| w(\boldsymbol{x},\boldsymbol{y}) \bigr| < \infty. \tag{2.5.20}$$

It is easy to see that $\widetilde{\rho}\circ\pi$ admits a continuous extension to the blow-up. Using (2.5.42), (2.5.18b) and (2.5.44) we deduce that for any $p>0$ there exists a constant $K_p>0$, such that

$$|x-y|^{m-1}\bigl| \Delta(\boldsymbol{x},\boldsymbol{y}) \bigr| \leq K_p\bigl(1+|x-y| \bigr)^{-p+m-1}, \ \ \forall \boldsymbol{x}\neq\boldsymbol{y} \tag{2.5.21}$$

Set

$$\delta(\xi,\eta) = \Delta\bigl(\pi(\xi,\eta) \bigr)$$

Since $\Delta(\boldsymbol{x},\boldsymbol{y})$ depends only on $\boldsymbol{y}-\boldsymbol{x}$ we deduce that $\delta(\xi,\eta)$ is independent of ξ. We have

$$I(R) = \int_{\mathfrak{X}} \Delta(\boldsymbol{x},\boldsymbol{y}) f_R^{\boxtimes 2}(\boldsymbol{x},\boldsymbol{y}) d\boldsymbol{x}d\boldsymbol{y}$$

$$= \frac{1}{2^{2m}}\int_{\mathbb{R}^m}\int_{|\boldsymbol{\nu}|=1,\, r\in(0,\infty)} r^{m-1}\delta\bigl(\xi, r\boldsymbol{\nu}\bigr) f_R\Bigl(\frac{\xi+r\boldsymbol{\nu}}{2} \Bigr) f_R\Bigl(\frac{\xi-r\boldsymbol{\nu}}{2} \Bigr) dr\, \mathrm{vol}_{S^{m-1}}[d\boldsymbol{\nu}]d\xi$$

$(\xi=2R\zeta)$

$$\stackrel{(2.5.43)}{=} R^m \underbrace{\int_{|\zeta|\leq 2r_0} 2^{-m}\left(\int_{\substack{|\boldsymbol{\nu}|=1\\ r>0}} r^{m-1}\delta\bigl(0,r\boldsymbol{\nu}\bigr) f\Bigl(\zeta+\frac{r\boldsymbol{\nu}}{2R} \Bigr) f(\zeta-\frac{r\boldsymbol{\nu}}{2R}\Bigr) dr\,\mathrm{vol}_{S^{m-1}}[d\boldsymbol{\nu}] \right) d\zeta}_{=:J(R)}.$$

We deduce from (2.5.43) and (2.5.44) that for any $p>0$ there exists $K_p>0$ such that for any $R>0$, $|\boldsymbol{\nu}|=1$ and $r>0$ we have

$$\left| r^{m-1}\delta\bigl(0,r\boldsymbol{\nu}\bigr) f\Bigl(\zeta+\frac{r\boldsymbol{\nu}}{2R} \Bigr) f\Bigl(\zeta-\frac{r\boldsymbol{\nu}}{2R} \Bigr) \right| \leq K_p\|f\|^2\bigl(1+r \bigr)^{-p+m-1}.$$

For $p>m$ we have

$$\int_{|\zeta|\leq 2r_0}\left(\int_{(0,\infty\times S^{m-1}} \bigl(1+r \bigr)^{-p+m-1} dr\,\mathrm{vol}_{S^{m-1}}[d\boldsymbol{\nu}] \right) d\zeta < \infty.$$

The dominated convergence theorem implies that $J(R)$ has a finite limit as $R\to\infty$. More precisely

$$\lim_{R\to\infty} J(R) = \int_{|\zeta|\leq 2r_0} \underbrace{\left(2^{-m}\int_{\substack{|\boldsymbol{\nu}|=1\\ r>0}} r^{m-1}\delta\bigl(0,r\boldsymbol{\nu}\bigr) dr\,\mathrm{vol}_{S^{m-1}}[d\boldsymbol{\nu}] \right)}_{=:Z_m(\mathfrak{a})} f(\zeta)^2 d\zeta.$$

This concludes the proof of Theorem 2.5.6 (ii) with $V_m(\mathfrak{a}) = Z_m(\mathfrak{a}) + C_m(\mathfrak{a})$.

□

As detailed in Appendix C.2 the above result can be rephrased as saying that the random measures $\mathcal{L}_N$ converge a.s. and L^2 to the deterministic measure $C_m(\mathfrak{a})\boldsymbol{\lambda}$. In particular, for any bounded Borel set S, the random variables $\mathcal{L}_N[S]$ converge a.s. to $C_m(\mathfrak{a})\boldsymbol{\lambda}[S]$ a.s.. Thus, in the white noise limit $N\to\infty$, the critical points of $\Phi_{\mathfrak{a}}^N$ tend to equidistribute with high confidence.

For any $f\in C^0_{\rm cpt}(\mathbb{R}^m)$ we have

$$\mathfrak{C}^N_{\mathfrak{a}}[f]=\mathfrak{C}_{\mathfrak{a}}[f_N],\;\; f_N(\boldsymbol{x}):=f\big(N^{-1}\boldsymbol{x}\big).$$

Suppose that $f\in C^0_{\rm cpt}(\mathbb{R}^n)$ is nonnegative and

$$\boldsymbol{\lambda}[f]=\int_{\mathbb{R}^m}f(\boldsymbol{x})d\boldsymbol{x}>0.$$

Then the sequence

$$f_N(\boldsymbol{x})=\frac{1}{N^m}f_N(\boldsymbol{x})$$

is asymptotically stationary in the precise sense defined in Appendix C.2. The random measure $\mathfrak{C}_{\mathfrak{a}}$ is stationary and Theorem C.2.2 implies that

$$\lim_{N\to\infty}\frac{1}{\lambda\big[\,f\,\big]}\mathfrak{C}^N_{\mathfrak{a}}[f]$$

exists a.s. and L^1 and it is a random variable $\widehat{\mathfrak{C}}_{\mathfrak{a}}$ independent of f. It called the *asymptotic intensity*. Moreover, we can identify $\widehat{\mathfrak{C}}_{\mathfrak{a}}$ with a measurable function on the space $\mathcal{M}$ of locally finite Borel measures on $\mathbb{R}^m$. In the sutuation at hand, the above results show that the random variable $\widehat{\mathfrak{C}}_{\mathfrak{a}}$ is deterministic, more precisely, $\widehat{\mathfrak{C}}_{\mathfrak{a}}=C_m(\mathfrak{a})$

The distribution of the random measure $\mathfrak{C}_{\mathfrak{a}}$ is a Borel probability measure $\mathbb{P}_{\mathfrak{C}_{\mathfrak{a}}}$ on $\mathcal{M}$. The additive group $\mathbb{R}^m$ acts on $\mathcal{M}$ by translations and, since $\Phi_{\mathfrak{a}}$ is stationary, we deduce that $\mathbb{P}_{\mathfrak{C}_{\mathfrak{a}}}$ is invariant under the above action of $\mathbb{R}^m$. The distribution of the random measure $\mathfrak{C}_{\mathfrak{a}}$ is a Borel probability measure $\mathbb{P}_{\mathfrak{C}_{\mathfrak{a}}}$ on $\mathcal{M}$. Since $\Phi_{\mathfrak{a}}$ is stationary, we deduce that $\mathbb{P}_{\mathfrak{C}_{\mathfrak{a}}}$ is invariant under the above action of $\mathbb{R}^m$.

The asymptotic intensity $\widehat{\mathfrak{C}}_{\mathfrak{a}}$ has an alternate, ergodic, description (C.2.2) involving the above action of $\mathbb{R}^m$ on $\mathcal{M}$. The fact that $\widehat{\mathfrak{C}}_{\mathfrak{a}}$ is in fact deterministic suggests that the action of $\mathbb{R}^m$ on $\big(\mathcal{M},\mathbb{P}_{\mathfrak{C}_{\mathfrak{a}}}\big)$ might be ergodic. There is some circumstantial evidence.

The Gaussian function $\Phi_{\mathfrak{a}}$ defines a Gaussian measure $\boldsymbol{\Gamma}$ on $C^2(\mathbb{R}^m)$. The additive group $\mathbb{R}^m$ acts on $C^2(\mathbb{R}^m)$ by translations. Since the Gaussian function $\Phi_{\mathfrak{a}}$ is stationary, we deduce that the above action is $\boldsymbol{\Gamma}$-preserving. Since the spectral measure of $\Phi_{\mathfrak{a}}$ is absolutely continuous with respect to the Lebesgue measure, the above action of $\mathbb{R}^m$ on $\big(C^2(\mathbb{R}^m),\boldsymbol{\Gamma}\big)$ is ergodic; see [**20**]. This is the fact used in 1960 by V. Volkovski [**156**] to prove Corollary 2.5.7 in the case $m=1$. We refer to [**37**, Sec. 11.5] for details.

2.5.3. Critical points of random Fourier series. Fix an amplitude $\mathfrak{a}\in\mathcal{S}(\mathbb{R})$. This means that $\mathfrak{a}$ is even and satisfies $\mathfrak{a}(0)=1$. Consider the random random Fourier series $F^R_{\mathfrak{a}}:\mathbb{R}^m\to\mathbb{R}$ defined by (1.2.21),

$$F^R_{\mathfrak{a}}(\vec{\theta})=R^{-m/2}\Big(\,A_0u_0+\sum_{\vec{\ell}\succ 0}\mathfrak{a}\big(\,\big|\,2\pi\vec{\ell}\,\big|/R\big)\big(\,A_{\vec{\ell}}u_{\vec{\ell}}(\vec{\theta})+B_{\vec{\ell}}v_{\vec{\ell}}(\vec{\theta})\,\big)\,\Big)$$

$$=R^{-m/2}\sum_{\vec{\ell}\in\mathbb{Z}^m}\mathfrak{a}\big(\,\big|\,2\pi\vec{\ell}\,\big|/R\,\big)Z_{\vec{\ell}}e^{2\pi\boldsymbol{i}\langle\vec{\ell},\theta\rangle},$$

where we recall that $(A_{\vec{\ell}})_{\vec{\ell}\in\mathbb{Z}^m}$, $(B_{\vec{\ell}})_{\vec{\ell}\in\mathbb{Z}^m}$ are i.i.d. standard normal variables

$$u_{\vec{\ell}}(\vec{\theta})=\sqrt{2}\cos 2\pi\langle\vec{\ell},\vec{\theta}\rangle,\;\; v_{\vec{\ell}}=\sqrt{2}\sin 2\pi\langle\vec{\ell},\vec{\theta}\rangle.$$

and $\succ$ is the lexicographic order on $\mathbb{R}^m$: $\boldsymbol{x} \prec \boldsymbol{y}$ iff there exists j such that $x_j < y_j$ and $x_i = y_i$, $\forall i < j$. The complex Gaussian variables $(Z_{\vec{\ell}})_{\vec{\ell}\in\mathbb{Z}^m}$ are defined as in (1.2.22). As explained in Example 1.2.31, the random function $R^{m/2}F^R_{\mathfrak{a}}$ converges in distribution to the Gaussian noise on the flat torus. For this reason we will refer to the $R \to \infty$ limist as white noise limits.

The covariance kernel of $F^R_{\mathfrak{a}}$ is $\mathcal{C}^R_{\mathfrak{a}}(\vec{\theta}, \vec{\varphi}) = C^R_{\mathfrak{a}}(\vec{\tau})$, where $\vec{\tau} = \vec{\theta} - \vec{\varphi}$. We deduce from (1.2.27) that

$$C^R_{\mathfrak{a}}(\vec{\tau}) = R^{-m} \sum_{\vec{\ell}\in\mathbb{Z}^m} \mathfrak{a}\big(\,|\,2\pi\vec{\ell}\,|/R\,\big)^2 e^{-2\pi \boldsymbol{i}\langle\vec{\ell},\vec{\tau}\rangle} \stackrel{(1.2.25)}{=} \sum_{\vec{k}\in\mathbb{Z}^m} \boldsymbol{K}_{\mathfrak{a}}\big(\,(\vec{k}-\vec{\tau})R\,\big), \tag{2.5.22}$$

and

$$\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle\xi,\boldsymbol{x}\rangle}\mathfrak{a}\big(\,|\xi|\,\big)^2\boldsymbol{\lambda}\big[\,d\xi\,\big].$$

We can interpret $F^R_{\mathfrak{a}}$ in two ways, either as a $\mathbb{Z}^m$-periodic stationary Gaussian function on $\mathbb{R}^m$ or as a Gaussian random function on the torus $\mathbb{T}^m/\mathbb{Z}^m$.

We set

$$\Psi^R_{\mathfrak{a}}(\boldsymbol{x}) := F^R_{\mathfrak{a}}(x/R),$$

and we think of $\Psi^R_{\mathfrak{a}}$ as a $(R\mathbb{Z})^m$-periodic random function on $\mathbb{R}^m$.

This is a stationary random function with covariance kernel $\mathcal{K}^R_{\mathfrak{a}}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{K}^R_{\mathfrak{a}}(\boldsymbol{x}-\boldsymbol{y})$, where

$$\boldsymbol{K}^R_{\mathfrak{a}}(\boldsymbol{x}) = C^R_{\mathfrak{a}}(\boldsymbol{x}/R) \stackrel{(2.5.22)}{=} \sum_{\vec{k}\in\mathbb{Z}^m} \boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x} - R\vec{k}).$$

Since $\boldsymbol{K}_{\mathfrak{a}}$ is a Schwartz function we deduce that

$$\lim_{R\to\infty} \boldsymbol{K}^R_{\mathfrak{a}}(\boldsymbol{x}) = \boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}),\ \ \forall \boldsymbol{x}\in\mathbb{R}^m,$$

so that $\Psi^R_{\mathfrak{a}}$ converges to in distribution to the isotropic function $\Phi_{\mathfrak{a}}$ as $R\to\infty$.

Clearly, $\nabla F^R_{\mathfrak{a}}(\boldsymbol{y}) = 0$ iff $\nabla\Psi^R_{\mathfrak{a}}(R\boldsymbol{y}) = 0$ so, for any box $B = [a,b]^m \subset \mathbb{R}^m$, and any $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$, we have

$$\mathfrak{C}\big[\,B, F^R_{\mathfrak{a}}\,\big] = \mathfrak{C}\big[\,RB, \Psi^R_{\mathfrak{a}}\,\big],\ \ \mathfrak{C}\big[\,f, F^R_{\mathfrak{a}}\,\big] = \mathfrak{C}\big[\,f_R, \Psi^R_{\mathfrak{a}}\,\big],$$

where $f_R(\boldsymbol{x}) = f\big(\,R^{-1}\boldsymbol{x}\,\big)$. In Example 2.3.8 we proved the equality (2.3.30)

$$\mathbb{E}\big[\,\mathfrak{C}[B, F^R_{\mathfrak{a}}]\,\big] = R^m\big(\,C_m(\mathfrak{a})\,\mathrm{vol}\,\big[\,B\,\big] + O\big(\,R^{-\infty}\,\big)\,\big),$$

and (2.3.31),

$$\mathfrak{C}\big[\,f, F^R_{\mathfrak{a}}\,\big] = R^m\Big(\,C_m(\mathfrak{a}) + O\big(\,R^{-\infty}\,\big)\,\Big)\int_{\mathbb{R}^m} f(\boldsymbol{x})d\boldsymbol{x},$$

where $C_m(\mathfrak{a})$ is defined in (2.3.26). In [**115**] we proved that that there exists a constant $C'_{\mathfrak{a}}(m) \geq 0$ such that if $B = [0,r]^m$ $r\in(0,1/2)$ we have

$$\lim_{R\to\infty} R^{-m}\,\mathrm{Var}\,\big[\,\mathfrak{C}[B, F^R_{\mathfrak{a}}]\,\big] = C'_m(\mathfrak{a})\bar{\lambda}\big[\,B\,\big]. \tag{2.5.23}$$

The proof of (2.5.23) in [**115**] is very laborious and computationally intensive.

The first result of this subsection is a functional version of (2.5.23). We achieve this using a less computational, more robust and more conceptual technique. One consequence of this asymptotic estimate is a (functional) strong law of large numbers concerning the random measures $\mathfrak{C}[-, F_{\mathfrak{a}}^N]$, $N \in \mathbb{N}$.

First some notation. Denote by $|-|$ the Euclidean norm on $\mathbb{R}^m$ and by $|-|_\infty$ the sup-norm on $\mathbb{R}^m$. For $\boldsymbol{x}_0 \in \mathbb{R}^m$ and $r > 0$ we set

$$B_r(x_0) := \{ \boldsymbol{x} \in \mathbb{R}^m;\ \ |\boldsymbol{x}| \leq r \},\ \ B_r^\infty(x_0) := \{ \boldsymbol{x} \in \mathbb{R}^m;\ \ |\boldsymbol{x}|_\infty \leq r \}.$$

Clearly $B_r(x_0) \subset B_r^\infty(x_0)$.

The function $F_{\mathfrak{a}}^R$ is $\mathbb{Z}^m$-periodic and for $r \in (0, 1/2)$ the ball $B_r^\infty(0)$ is contained in the interior of a fundamental domain of the $\mathbb{Z}^m$-action since $|\boldsymbol{x} - \boldsymbol{y}|_\infty \leq 2r < 1$ and $|\vec{\ell}|_\infty \geq 1$, $\forall \vec{\ell} \in \mathbb{Z}^m \setminus 0$. This reflects the fact that the injectivity radius of the flat torus $\mathbb{T}^m = \mathbb{R}^m/\mathbb{Z}^m$ is $\leq \frac{1}{2}$ so $B_r(0)$ can be viewed as a geodesic ball. We can now state the main technical result of this paper.

Theorem 2.5.9. *Fix an amplitude $\mathfrak{a}$, a positive integer $m \in \mathbb{N}$ and a radius $r_0 \in (0, 1/2)$. There exists a constant $V = V_m(\mathfrak{a}) \geq 0$ that depends only on m and $\mathfrak{a}$ such that, for any continuous function $f : \mathbb{R}^m \to \mathbb{R}$ with support contained in $B_{r_0}(0)$, we have*

$$\lim_{R\to\infty} R^{-m} \operatorname{Var}\big[\, \mathfrak{C}[f, F_{\mathfrak{a}}^R] \,\big] = V_m(\mathfrak{a}) \int_{\mathbb{R}^m} f(\boldsymbol{x})^2 d\boldsymbol{x}. \tag{2.5.24}$$

Before we present the proof of this theorem let us discuss some of its consequences. Consider the normalized random measures

$$\bar{\mathfrak{C}}_R := \frac{1}{R^m} \mathfrak{C}[-, F_{\mathfrak{a}}^R],\ \ R > 0.$$

A function then we deduce that for any $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$, $\operatorname{supp} f \in B_{r_0}(0)$, we have

$$\lim_{R\to\infty} \mathbb{E}\big[\, \bar{\mathfrak{C}}_R[f] \,\big] = C_m(\mathfrak{a}) \int_{\mathbb{R}^m} f(\boldsymbol{x}) \boldsymbol{\lambda}\big[\, dx \,\big], \tag{2.5.25}$$

and

$$\operatorname{Var}\big[\, \bar{\mathfrak{C}}_R[f] \,\big] \sim V_m(\mathfrak{a}) R^{-m} \int_{\mathbb{R}^m} f(\boldsymbol{x})^2 d\boldsymbol{x} \ \text{ as } R \to \infty. \tag{2.5.26}$$

Since $F_{\mathfrak{a}}^R$ is stationary, the same is true for any continuous function f with support contained in $B_{r_0}(\boldsymbol{x}_0)$. Indeed, this follows by applying (2.5.26) to the function $\tilde{f}(\boldsymbol{x}) = f(\boldsymbol{x} - \boldsymbol{x}_0)$.

Using finite partitions of unity we can represent any $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$ as a finite sum of functions supported in Euclidean balls of radius r_0 and we deduce from (2.5.24) that

$$\forall f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m),\ \ \operatorname{Var}\big[\, \bar{\mathfrak{C}}_R[f] \,\big] = O\big(\, R^{-m} \,\big) \ \text{ as } R \to \infty.$$

If $m \geq 2$, then

$$\sum_{N\in\mathbb{N}} \frac{1}{N^m} < \infty$$

The Borel-Cantelli lemma and (2.5.26) imply that for any nonnegative $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$ we have

$$\lim_{N\to\infty} \bar{\mathfrak{C}}_N[f] = C_m(\mathfrak{a}) \int_{\mathbb{R}^m} f(\boldsymbol{x}) \boldsymbol{\lambda}\big[\, dx \,\big] \ \text{ a.s. and in } L^2. \tag{2.5.27}$$

Thus, in the white noise limit ($R \to \infty$), the critical points of $F_{\mathfrak{a}}^R$ will equidistribute with probability 1. In the case $m = 1$, this law of large numbers is proved in the recent work of L. Gass [**65**, Thm. 1.6].

We can rephrase the equality (2.5.27) as a law as large numbers. We refer to Appendix C.2 for the various concepts of convergence of random measures. The equality (2.5.27) and Theorem C.2.1 imply the following result.

Corollary 2.5.10 (Strong Law of Large Numbers). *In white noise limit ($N \to \infty$) the* random *measures* $\frac{1}{N^m}\mathfrak{C}\big[-, F_{\mathfrak{a}}^N\big]$ *converge vaguely a.s. and* L^2 *to the* deterministic *measure* $C_m(\mathfrak{a})\boldsymbol{\lambda}$. *In particular, for any bounded Borel subset* $S \subset \mathbb{R}^m$ *we have*

$$\lim_{N\to\infty} \frac{1}{N^m}\mathfrak{C}\big[\,S, F_{\mathfrak{a}}^N\,\big] = C_m(\mathfrak{a})\boldsymbol{\lambda}\big[\,S\,\big]$$

a.s. and L^2. □

In [**116**] we proved that in the white noise limit the *random* measures $\mathfrak{C}\big[-, F_{\mathfrak{a}}^N\big]$ also satisfy a Central Limit Theorem. More precisely, for any $r \in (0,1)$

$$\frac{1}{N^{m/2}}\Big(\,\mathfrak{C}[B^\infty_{r/2}, F_{\mathfrak{a}}^N] - \mathbb{E}\big[\,\mathfrak{C}[B^\infty_{r/2}, F_{\mathfrak{a}}^N]\,\big]\,\Big)$$

converges in distribution to a centered normal random variable with nonzero variance.

As explained in Appendix C.2, each stationary random measure $\mathfrak{M}$ on $\mathbb{R}^m$ has an asymptotic intensity $\widehat{\mathfrak{M}}$. This is a random variable defined by the ergodic limit (C.2.3)

$$\lim_{N\to\infty}\frac{1}{\operatorname{vol}[NC]}\mathfrak{M}\big[\,NC\,\big] = \widehat{\mathfrak{M}}, \tag{2.5.28}$$

where $C \subset \mathbb{R}^m$ is any compact convex subset of $\mathbb{R}^m$ containing the origin in the interior.

The random measure $\mathfrak{M} = \mathfrak{C}[-, \Phi_{\mathfrak{a}}]$ is stationary and the results of the previous subsection show that the asymptotic intensity of $\mathfrak{C}[-, \Phi_{\mathfrak{a}}]$ is the constant $\widehat{\mathfrak{C}}_{\Phi_{\mathfrak{a}}} = C_m(\mathfrak{a})$.

For fixed $N > 0$, the random function $\Phi_{\mathfrak{a}}^{N_0}$ is $\big(R\mathbb{Z}\big)^m$-periodic and we deduce that for any $N \in \mathbb{N}$ we have

$$\mathfrak{C}[NB^\infty_{R/2}, \Phi_{\mathfrak{a}}^R] = N^m\mathfrak{C}[B^\infty_{R/2}, \Phi_{\mathfrak{a}}^R].$$

Hence

$$\mathfrak{C}[B^\infty_{R/2}, \Phi_{\mathfrak{a}}^R] = \lim_{N\to\infty}\frac{1}{N^m}\mathfrak{C}[NB^\infty_{R/2}, \Phi_{\mathfrak{a}}^R] \overset{(2.5.28)}{=} \widehat{\mathfrak{C}}_{\Phi_{\mathfrak{a}}^R}\operatorname{vol}\big[\,B^\infty_{R/2}\,\big],$$

where $\widehat{\mathfrak{C}}_{\Phi_{\mathfrak{a}}^R}$ denotes the asymptotic intensity of the stationary random measure $\mathfrak{C}[-, \Phi_{\mathfrak{a}}^{N_0}]$. Hence

$$\widehat{\mathfrak{C}}_{\Phi_{\mathfrak{a}}^R} = \frac{1}{R^m}\mathfrak{C}\big[\,B^\infty_{R/2}, \Phi_{\mathfrak{a}}^R\,\big] = \frac{1}{\operatorname{vol}[B^\infty_{R/2}]}\mathfrak{C}\big[\,B^\infty_{R/2}, \Phi_{\mathfrak{a}}^R\,\big].$$

Corollary 2.5.10 shows that

$$\lim_{N\to\infty}\widehat{\mathfrak{C}}_{\Phi_{\mathfrak{a}}^N} = \widehat{\mathfrak{C}}_{\Phi_{\mathfrak{a}}} = C_m(\mathfrak{a}),$$

a.s. and L^2.

Proof of Theorem 2.5.9. We follow the strategy in [**62**]. We split the proof of Theorem 2.5.9 into several conceptually distinct parts.

1. The key estimate. The following technical result will play a key role.

Lemma 2.5.11. *Fix a box* $B = B^\infty_{r_0/2}(0) = [-r_0/2, r_0/2]^m$, $r_0 \in (0,1)$. *Then the following hold.*

(i) *For any $\ell \in \mathbb{N}_0$ and any $p > m$ there exists $C = C(p, m, \ell, \mathfrak{a}) > 0$, independent of R, such that, $\forall R > 2$*

$$\| \boldsymbol{K}_{\mathfrak{a}}^R - \boldsymbol{K}_{\mathfrak{a}} \|_{C^\ell(RB)} \leq CR^{-p}$$

(ii) *For any $\ell \in \mathbb{N}_0$ and any $p > m$ there exists $C = C(p, m, \ell, \mathfrak{a}) > 0$, independent of R, such that, $\forall R > 2$, $\forall \boldsymbol{x}, \boldsymbol{y} \in RB$*

$$\left| D^\ell \boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{x} - \boldsymbol{y}) \right| \leq \frac{C}{\left(1 + |\boldsymbol{x} - \boldsymbol{y}|_\infty \right)^p}.$$

Proof. (i) Denote by $\mathcal{T}_{R\vec{k}} \boldsymbol{K}_{\mathfrak{a}}$ the translate

$$\mathcal{T}_{R\vec{k}} \boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}) := \boldsymbol{K}\left(\boldsymbol{x} - R\vec{k} \right).$$

We have

$$\boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{x}) - \boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x}) = \sum_{\vec{k} \in \mathbb{Z}^m \setminus 0} \mathcal{T}_{R\vec{k}} \boldsymbol{K}_{\mathfrak{a}}\left(\boldsymbol{x} \right).$$

Now observe that $\forall R > 0$, $\forall \boldsymbol{x} \in RB$, and any $\vec{k} \in \mathbb{Z}^m \setminus 0$ we have

$$\left| \boldsymbol{x} - R\vec{k} \right|_\infty \geq N \left| \vec{k} \right|_\infty - \left| \boldsymbol{x} \right|_\infty \geq R\left(\left| \vec{k} \right|_\infty - r_0/2 \right).$$

Since $\boldsymbol{K}_{\mathfrak{a}}$ and all its derivatives are Schwartz functions we deduce that for any $p > m$, and any $\vec{k} \in \mathbb{Z}^m \setminus 0$

$$\left\| \mathcal{T}_{R\vec{k}} \boldsymbol{K}_{\mathfrak{a}} \right\|_{C^\ell(NB)} \leq C(p, m, \ell, \mathfrak{a}) R^{-p} \left(\left| \vec{k} \right|_\infty - r_0/2 \right)^{-p}.$$

The last expression is well defined since $r < 1 \leq \left| \vec{k} \right|_\infty$ for any $\vec{k} \in \mathbb{Z}^m \setminus 0$. Hence

$$\left\| \boldsymbol{K}_{\mathfrak{a}}^R - \boldsymbol{K}_{\mathfrak{a}} \right\|_{C^\ell(NB)} \leq C(p, m, \ell, \mathfrak{a}) R^{-p} \sum_{\vec{k} \in \mathbb{Z}^m \setminus 0} \left(\left| \vec{k} \right|_\infty - r_0/2 \right)^{-p}$$

The above series is convergent since $p > m$.

(ii) Note that $\forall \boldsymbol{x}, \boldsymbol{y} \in RB$ we have $\left| \boldsymbol{x} - \boldsymbol{y} \right|_\infty \leq Rr_0$. Set $\boldsymbol{z} := \boldsymbol{x} - \boldsymbol{y}$. We discuss only the case $\ell = 0$. The general case can be reduced to this case by taking partial derivatives.

Using (i) we deduce that

$$C = \sup_R \sup_{\left| \boldsymbol{z} \right|_\infty < r_0} \left| \boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{z}) \right| < \infty$$

and thus, $\forall R \geq 2$, $\forall \left| \boldsymbol{z} \right|_\infty < r_0$,

$$\left| \boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{z}) \right| < \frac{C\left(1 + r_0 \right)^p}{\left(1 + \left| \boldsymbol{z} \right|_\infty \right)^p}.$$

Assume now that $\left| \boldsymbol{z} \right|_\infty \geq r_0$. We have

$$\boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{z}) = \boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{z}) + \sum_{\vec{k} \in \mathbb{Z}^m \setminus 0} \mathcal{T}_{R\vec{k}} \boldsymbol{K}_{\mathfrak{a}}\left(\boldsymbol{z} \right),$$

and thus,

$$\left| \boldsymbol{K}_{\mathfrak{a}}^R(\boldsymbol{z}) \right| \leq \left| \boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{z}) \right| + \sum_{\vec{k} \in \mathbb{Z}^m \setminus 0} \left| \mathcal{T}_{R\vec{k}} \boldsymbol{K}_{\mathfrak{a}}\left(\boldsymbol{z} \right) \right|.$$

Since $\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{x})$ is Schwartz we deduce that there exists a constant $C = C(p, \mathfrak{a})$ such that

$$\frac{C_p}{\left(1 + |\,\boldsymbol{z}\,|_\infty\right)^p} + C_p \sum_{\vec{k} \in \mathbb{Z}^m \setminus 0} \frac{1}{\left(1 + |\,\boldsymbol{z} - R\vec{k}\,|_\infty\right)^p}$$

We have $|\,\boldsymbol{z}\,|_\infty \leq Rr_0$ and

$$|\,\boldsymbol{z} - Z\vec{k}\,|_\infty \geq |\,\boldsymbol{z}\,|_\infty \Big(\frac{R|\,\vec{k}\,|_\infty}{|\,\boldsymbol{z}\,|_\infty} - 1 \Big) \geq |\,\boldsymbol{z}\,|_\infty \Big(\frac{1}{r_0} |\,\vec{k}\,|_\infty - 1 \Big)$$

Thus

$$\sum_{\vec{k} \in \mathbb{Z}^m \setminus 0} \frac{1}{\left(1 + |\,\boldsymbol{z} - R\vec{k}\,|_\infty\right)^p} \leq |\,\boldsymbol{z}\,|_\infty^{-p} \underbrace{\sum_{\vec{k} \in \mathbb{Z}^m \setminus 0} \Big(\frac{1}{r_0} |\,\vec{k}\,|_\infty - 1 \Big)^{-p}}_{<\infty}.$$

□

2. An integral formula. Set

$$B := B^\infty_{r_0/2}(0), \;\; f_R(\boldsymbol{x}) := f(\boldsymbol{x}/R),$$

$$Z^R[f] := \mathfrak{C}[f, F^R_{\mathfrak{a}}] = \mathfrak{C}[f_R, \Phi^R_{\mathfrak{a}}], \;\; Z[f] := \mathfrak{C}[f, \Phi_{\mathfrak{a}}].$$

Denote by $\rho^R_{\mathfrak{a}}$ the Kac-Rice density of $\nabla\Phi^R_{\mathfrak{a}}$ and by $\rho_{\mathfrak{a}}$ the Kac-Rice density of $\nabla\Phi_{\mathfrak{a}}$; see (2.2.10). Since both $\Phi^R_{\mathfrak{a}}$ and $\Phi_{\mathfrak{a}}$ are stationary random functions we deduce that both $\rho^R_{\mathfrak{a}}$ and $\rho_{\mathfrak{a}}$ are constant functions.

The covariance functions $\boldsymbol{K}^R_{\mathfrak{a}}(\boldsymbol{z})$ and $\boldsymbol{K}_{\mathfrak{a}}(\boldsymbol{z})$ are even, so the odd order derivatives of these functions vanish at 0. This implies that the Gaussian vectors $\mathrm{Hess}_{\Phi^R_{\mathfrak{a}}}(0)$ and $\nabla\Phi^R_{\mathfrak{a}}(0)$ are independent. A similar phenomenon is true for $\Phi_{\mathfrak{a}}$. Thus, the conditional expectations in the Kac-Rice formula are usual expectations. Using Lemma 1.1.27 and Lemma 2.5.11(i) we deduce that for any $\boldsymbol{x} \in \mathbb{R}^m$

$$\sup_{\boldsymbol{x} \in RB} \big|\, \rho^R_{\mathfrak{a}}(\boldsymbol{x}) - \rho_{\mathfrak{a}}(\boldsymbol{x}) \,\big| = \big|\, \rho^R_{\mathfrak{a}}(0) - \rho_{\mathfrak{a}}(0) \,\big| = O\big(\, R^{-\infty} \,\big), \tag{2.5.29}$$

where $O(R^{-\infty})$ is short-hand for $O(R^{-p})$, $\forall p > 0$ as $R \to \infty$. We deduce that

$$R^{-m}\big(\mathbb{E}\big[\, Z^R[f] \,\big] - \mathbb{E}\big[\, Z[f] \,\big] \big) = R^{-m} \int_{RB} f_R(\boldsymbol{x}) \big(\, \rho^R_{\mathfrak{a}}(0) - \rho_{\mathfrak{a}}(0) \,\big) d\boldsymbol{x} =$$

$$= \int_B f(\boldsymbol{y}) \big(\, \rho^R_{\mathfrak{a}}(0) - \rho_{\mathfrak{a}}(0) \,\big) d\boldsymbol{y} = O\big(\, R^{-\infty} \,\big).$$

We need to introduce some additional notation.

- $\Phi^\infty_{\mathfrak{a}} = \Phi_{\mathfrak{a}}$.
- For any $R \in (0, \infty]$ we define

$$\widehat{\Phi}^R, \; \widehat{\Phi} : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R},$$

$$\widehat{\Phi}^R(\boldsymbol{x}, \boldsymbol{y}) = \Phi^R_{\mathfrak{a}}(\boldsymbol{x}) + \Phi^R_{\mathfrak{a}}(\boldsymbol{y}), \;\; \widehat{\Phi}(\boldsymbol{x}, \boldsymbol{y}) = \Phi_{\mathfrak{a}}(\boldsymbol{x}) + \Phi_{\mathfrak{a}}(\boldsymbol{y}),$$

$$\hat{\mathfrak{C}}^R := \mathfrak{C}[-, \widehat{\Phi}^R_{\mathfrak{a}}], \;\; \widehat{H}_R(\boldsymbol{x}, \boldsymbol{y}) := \mathrm{Hess}_{\widehat{\Phi}^R}(\boldsymbol{x}, \boldsymbol{y}), \;\; H_R(\boldsymbol{x}) := \mathrm{Hess}_{\Phi^R_{\mathfrak{a}}}(\boldsymbol{x}).$$

- Choose an independent copy $\Psi_{\mathfrak{a}}^{R}$ of $\Phi_{\mathfrak{a}}^{R}$ and for $R \in (0, \infty]$ set
$$\widetilde{\Phi}^{R}(\boldsymbol{x}, \boldsymbol{y}) := \Phi_{\mathfrak{a}}^{R}(\boldsymbol{x}) + \Psi_{\mathfrak{a}}^{R}(\boldsymbol{y}), \;\; \widetilde{H}_{R}(\boldsymbol{x}, \boldsymbol{y}) := \operatorname{Hess}_{\widetilde{\Phi}^{R}}(\boldsymbol{x}, \boldsymbol{y}),$$
$\tilde{\mathfrak{C}}^{R} = \mathfrak{C}[-, \widetilde{\Phi}^{N}]$.
- For $R \in (0, \infty)$ define
$$f_{R}^{\boxtimes 2} : \mathbb{R}^{m} \times \mathbb{R}^{m} \to \mathbb{R}, \;\; f_{R}^{\boxtimes}(\boldsymbol{x}, \boldsymbol{y}) = f_{R}(\boldsymbol{x}) f_{R}(\boldsymbol{y})$$
and set $\|f\| := \|f\|_{C^{0}(\mathbb{R}^{m})}$.
- Set
$$\mathfrak{X} = \mathbb{R}^{m} \times \mathbb{R}^{m} \setminus \Delta = \big\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{m} \times \mathbb{R}^{m}; \;\; \boldsymbol{x} \neq \boldsymbol{y} \big\}.$$

Observe that the random function on $\widetilde{\Phi}^{R}(\boldsymbol{x}, \boldsymbol{y})$ is *stationary* with respect to the action of $\mathbb{R}^{2m}$ on itself by translation, while $\widehat{\Phi}_{\mathfrak{a}}^{R}$ is stationary with respect to the diagonal action by translations of $\mathbb{R}^{m}$ on $\mathbb{R}^{m} \times \mathbb{R}^{m}$,
$$T_{\boldsymbol{v}}(\boldsymbol{x}, \boldsymbol{y}) = \big(\boldsymbol{v} + \boldsymbol{x}, \boldsymbol{v} + \boldsymbol{y} \big), \;\; \forall \boldsymbol{v}, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{m}.$$
We have
$$\hat{\mathfrak{C}}^{R}[\boldsymbol{I}_{\mathfrak{X}} f_{R}^{\boxtimes 2}] = \sum_{\substack{\nabla \Phi_{\mathfrak{a}}^{R}(\boldsymbol{x}) = \nabla \Phi_{\mathfrak{a}}^{R}(\boldsymbol{y}) = 0, \\ \boldsymbol{x} \neq \boldsymbol{y}}} f_{R}(\boldsymbol{x}) f_{R}(\boldsymbol{y}) = Z^{R}[f]^{2} - Z^{R}[f^{2}].$$
Bulinskaya's lemma implies that
$$\mathbb{P}\big[\exists \boldsymbol{x} : \;\; \nabla \Phi_{\mathfrak{a}}(\boldsymbol{x}) = \nabla \Psi_{\mathfrak{a}}(\boldsymbol{x}) = 0 \big] = 0$$
and we deduce
$$\tilde{\mathfrak{C}}^{R}[\boldsymbol{I}_{\mathfrak{X}} f_{R}^{\boxtimes 2}] = \sum_{\substack{\nabla \Phi_{\mathfrak{a}}^{R}(\boldsymbol{x}) = \nabla \Psi_{\mathfrak{a}}^{R}(\boldsymbol{y}) = 0, \\ \boldsymbol{x} \neq \boldsymbol{y}}} f_{R}(\boldsymbol{x}) f_{R}(\boldsymbol{y})$$
$$= \sum_{\nabla \Phi_{\mathfrak{a}}^{R}(\boldsymbol{x}) = \nabla \Psi_{\mathfrak{a}}^{R}(\boldsymbol{y}) = 0} f_{R}(\boldsymbol{x}) f_{R}(\boldsymbol{y}) = \mathfrak{C}[f, \Phi_{\mathfrak{a}}^{R}] \big] \mathfrak{C}[f, \Psi_{\mathfrak{a}}^{R}], \;\; \text{a.s..}$$
Hence
$$\mathbb{E}\big[\mathfrak{C}[f, \Phi_{\mathfrak{a}}^{R}] \mathfrak{C}[f, \Psi_{\mathfrak{a}}^{R}] \big] = \mathbb{E}\big[\mathfrak{C}[f, \Phi_{\mathfrak{a}}^{R}] \big] \cdot \mathbb{E}\big[\mathfrak{C}[f, \Psi_{\mathfrak{a}}^{R}] \big] = \mathbb{E}\big[\mathfrak{C}[f, \Phi_{\mathfrak{a}}^{R}] \big]^{2}$$
so that
$$\mathbb{E}\big[\hat{\mathfrak{C}}^{R}[\boldsymbol{I}_{\mathfrak{X}} f_{R}^{\boxtimes 2}] \big] - \mathbb{E}\big[\tilde{\mathfrak{C}}^{R}[\boldsymbol{I}_{\mathfrak{X}} f_{R}^{\boxtimes 2}] \big] = \underbrace{\mathbb{E}\big[Z^{R}[f]^{2} \big] - \mathbb{E}\big[Z^{R}[f] \big]^{2}}_{= \operatorname{Var}\big[Z^{R}[f] \big]} - \mathbb{E}\big[Z^{R}[f^{2}] \big]$$
We have seen that
$$\lim_{R \to \infty} R^{-m} \mathbb{E}\big[Z^{R}[f^{2}] \big] = C_{m}(\mathfrak{a}) \int_{\mathbb{R}^{m}} f^{2}(\boldsymbol{x}) d\boldsymbol{x}$$
so we have to show that
$$I(R) := \mathbb{E}\big[\hat{\mathfrak{C}}^{R}[\boldsymbol{I}_{\mathfrak{X}} f_{R}^{\boxtimes 2}] \big] - \mathbb{E}\big[\tilde{\mathfrak{C}}^{R}[\boldsymbol{I}_{\mathfrak{X}} f_{R}^{\boxtimes 2}] \big] \sim c R^{-m} \;\; \text{as } R \to \infty \tag{2.5.30}$$
for some constant $c \in \mathbb{R}$.

According to Corollary 2.4.24, there exists $R_{0} > 0$ such that for $R \geq R_{0}$, the gradient $\nabla \Phi_{\mathfrak{a}}^{R}$ is 2-ample and Φ_{R} is J_{1}-ample so, for $R \geq R_{0}$ the gradient $\nabla \widehat{\Phi}^{R}(\boldsymbol{x}, \boldsymbol{y})$ is nondegenerate for any $\boldsymbol{x} \neq \boldsymbol{y}$ and the random vector $\big(\Phi_{\mathfrak{a}}^{R}(\boldsymbol{x}), \nabla \Phi_{\mathfrak{a}}^{R} \big)$ is nondegenerate for any $\boldsymbol{x} \in \mathbb{R}^{n}$. As shown in Example 1.2.35 this is true also for $R = \infty$, where we recall that $\Phi_{\mathfrak{a}}^{\infty} = \Phi_{\mathfrak{a}}^{R}$.

We can apply the Kac-Rice formula and we deduce that for any $R > R_0$ we have

$$\begin{gathered}\mathbb{E}\big[\,\hat{\mathfrak{C}}^R[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}]\,\big]\\ = \int_{\mathbb{R}^m\times\mathbb{R}^m\setminus\Delta} \underbrace{\mathbb{E}\big[\,|\det \widehat{H}_R(\boldsymbol{x},\boldsymbol{y})|\,\big|\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y}) = 0\,\big] p_{\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})}(0)}_{=\widehat{\rho}_R(\boldsymbol{x},\boldsymbol{y})} f_R^{\boxtimes 2}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}d\boldsymbol{y}\,\big].\end{gathered} \tag{2.5.31}$$

The gradient $\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})$ is nondegenerate for any $\boldsymbol{x},\boldsymbol{y}$ and invoking Kac-Rice again we obtain

$$\begin{gathered}\mathbb{E}\big[\,\tilde{\mathfrak{C}}^R[\boldsymbol{I}_{\mathfrak{X}} f_R^{\boxtimes 2}]\,\big]\\ = \int_{\mathbb{R}^m\times\mathbb{R}^m\setminus\Delta} \underbrace{\mathbb{E}\big[\,|\det \widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})|\,\big|\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y}) = 0\,\big] p_{\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})}(0)}_{=\widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})} f_R^{\boxtimes 2}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}d\boldsymbol{y}\,\big].\end{gathered} \tag{2.5.32}$$

The function $\widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})$ is independent of $\boldsymbol{x},\boldsymbol{y}$ since the random function $\widetilde{\Phi}^R$ is stationary. Thus

$$\begin{aligned} I(R) &= \int_{\mathfrak{X}} \big(\,\widehat{\rho}_R(\boldsymbol{x},\boldsymbol{y}) - \widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})\,\big) f_R(\boldsymbol{x}) f_R(\boldsymbol{y})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}d\boldsymbol{y}\,\big]\\ &= \int_{\substack{|\boldsymbol{x}|,\,|\boldsymbol{y}|\leq Rr_0/2,\\ \boldsymbol{x}\neq\boldsymbol{y}}} \big(\,\widehat{\rho}_R(\boldsymbol{x},\boldsymbol{y}) - \widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})\,\big) f_R(\boldsymbol{x}) f_R(\boldsymbol{y})\boldsymbol{\lambda}\big[\,d\boldsymbol{x}d\boldsymbol{y}\,\big].\end{aligned} \tag{2.5.33}$$

Let us observe that for any $\boldsymbol{x}\neq\boldsymbol{y}$ we have

$$\lim_{R\to\infty}\big(\,\widehat{\rho}_R(\boldsymbol{x},\boldsymbol{y}) - \widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})\,\big) = \big(\,\widehat{\rho}_\infty(\boldsymbol{x},\boldsymbol{y}) - \widetilde{\rho}_\infty(\boldsymbol{x},\boldsymbol{y})\,\big).$$

Moreover

$$\lim_{R\to\infty} f_R(\boldsymbol{x}) = f(0)$$

uniformly on compacts.

3.Off-diagonal behavior. Note that

$$\operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix} \operatorname{Var}\big[\,H_R(\boldsymbol{x})\,\big] & 0\\ 0 & \operatorname{Var}\big[\,H_R(\boldsymbol{y})\,\big]\end{bmatrix}.$$

For every $\boldsymbol{z}\in\mathbb{R}^m$ we set

$$T_R(\boldsymbol{z}) := \sum_{|\alpha|\leq 4}\big|\,\partial^\alpha \boldsymbol{K}^R_{\mathfrak{a}}(\boldsymbol{z})\,\big|.$$

Lemma 2.5.11(ii) shows that for every $p > 0$ there exists $C_p = C_p(\mathfrak{a}, m, r) > 0$ such that, $\forall R$, $\forall\big|\,\boldsymbol{z}\,\big|_\infty < Nr$

$$\forall N,\ \ \forall\big|\,\boldsymbol{z}\,\big|_\infty < Rr_0,\ \ T_R(z) \leq C_p\big(\,1 + \big|\,\boldsymbol{z}\,\big|_\infty\,\big)^{-p}. \tag{2.5.34}$$

We want to emphasize that C_p *is independent of* R.

Observe next that

$$\operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix} \operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & 0\\ 0 & \operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big]\end{bmatrix},$$

is independent of $\boldsymbol{x}$ and $\boldsymbol{y}$.

$$\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix} \operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x}), \nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big]\\ \operatorname{Cov}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y}), \nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big]\end{bmatrix}$$

$$= \operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] + \underbrace{\begin{bmatrix} 0 & \operatorname{Cov}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big] \\ \operatorname{Cov}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & 0 \end{bmatrix}}_{=:\mathcal{E}^R_{\nabla}(\boldsymbol{x},\boldsymbol{y})}.$$

Hence

$$\big\|\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] - \operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]\big\|_{\mathrm{op}} = \|\mathcal{E}^R_{\nabla}(\boldsymbol{x},\boldsymbol{y})\|_{\mathrm{op}} = O\big(T_R(\boldsymbol{x}-\boldsymbol{y})\big), \tag{2.5.35}$$

where $\|-\|_{\mathrm{op}}$ denotes the operator norm. Above and in the sequel, the *constant implied by the Landau symbol O is independent of R as long as $\boldsymbol{x},\boldsymbol{y}\in RB$.* In particular

$$\begin{gathered}\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1} = \Big(\operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] + \mathcal{E}^R_{\nabla}(\boldsymbol{x},\boldsymbol{y})\Big)^{-1} \\ = \operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\Big(\mathbb{1} + \operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\mathcal{E}^R_{\nabla}(\boldsymbol{x},\boldsymbol{y})\Big)^{-1}.\end{gathered} \tag{2.5.36}$$

We have shown in (2.3.20) that there exists an explicit positive constant d_m such that

$$\operatorname{Var}\big[\,\nabla\Phi_{\mathfrak{a}}(\boldsymbol{x})\,\big] = d_m\mathbb{1}_m,\ \ \forall\boldsymbol{x}.$$

Then $\operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] = \operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(0)\,\big]$, $\forall\boldsymbol{x}\in\mathbb{R}^m$ and

$$\operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(0)\,\big] = d_m\mathbb{1}_m + O\big(R^{-\infty}\big).$$

The variance $\operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]$ is independent of $\boldsymbol{x}$ and $\boldsymbol{y}$ and

$$\operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] = \operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(0)\,\big]\oplus\operatorname{Var}\big[\,\nabla\Phi^R_{\mathfrak{a}}(0)\,\big] = d_m\mathbb{1}_{2m} + O\big(R^{-\infty}\big). \tag{2.5.37}$$

From (2.5.36) and (2.5.37) we conclude that there exists $C_0>0$, independent of $R>R_0$, such that

$$\|\operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\mathcal{E}^R_{\nabla}(\boldsymbol{x},\boldsymbol{y})\|_{\mathrm{op}} < \frac{1}{2},\ \ \forall\boldsymbol{x},\boldsymbol{y}\in RB,\ \ |\boldsymbol{x}-\boldsymbol{y}|_\infty > C_0,$$

and thus

$$\begin{gathered}\big\|\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1} - \operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\big\|_{\mathrm{op}} \\ = O\big(T_R(\boldsymbol{x}-\boldsymbol{y})\big),\ \ \forall\boldsymbol{x},\boldsymbol{y}\in RB,\ \ |\boldsymbol{x}-\boldsymbol{y}|_\infty > C_0.\end{gathered} \tag{2.5.38}$$

Note that since $\Phi^R_{\mathfrak{a}}$ is stationary, $\operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big]$ is *independent* of $\boldsymbol{x}$ and $\boldsymbol{y}$.

$$\operatorname{Var}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix}\operatorname{Var}\big[\,H_R(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,H_R(\boldsymbol{x}),H_R(\boldsymbol{y})\,\big] \\ \operatorname{Cov}\big[\,H_R(\boldsymbol{y}),H_R(\boldsymbol{x})\,\big] & \operatorname{Var}\big[\,H_R(\boldsymbol{y})\,\big]\end{bmatrix}$$

$$= \operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big] + \underbrace{\begin{bmatrix} 0 & \operatorname{Cov}\big[\,H_R(\boldsymbol{x}),H_R(\boldsymbol{y})\,\big] \\ \operatorname{Cov}\big[\,H_R(\boldsymbol{y}),H_R(\boldsymbol{x})\,\big] & 0\end{bmatrix}}_{=:\mathcal{E}^R_H(\boldsymbol{x},\boldsymbol{y})}.$$

We deduce

$$\big\|\operatorname{Var}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big] - \operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big]\big\|_{\mathrm{op}} = \|\mathcal{E}^R_H(\boldsymbol{x},\boldsymbol{y})\|_{\mathrm{op}} = O\big(T_R(\boldsymbol{x}-\boldsymbol{y})\big). \tag{2.5.39}$$

We denote by $\widetilde{H}_R(\boldsymbol{x},y)^\flat$ the Gaussian random matrix

$$\widetilde{H}_R(\boldsymbol{x},y)^\flat = \widetilde{H}_R(\boldsymbol{x},y) - \mathbb{E}\big[\,\widetilde{H}_R(\boldsymbol{x},y)\,\|\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big].$$

We define $\widehat{H}_R(\boldsymbol{x},y)^\flat$ similarly

$$\widehat{H}_R(\boldsymbol{x},y)^\flat = \widehat{H}_R(\boldsymbol{x},y) - \mathbb{E}\big[\,\widehat{H}_R(\boldsymbol{x},y)\,\|\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big].$$

The distributions of $\widetilde{H}_R(\boldsymbol{x},y)^\flat$ and $\widehat{H}_R(\boldsymbol{x},y)^\flat$ are determined by the Gaussian regression formula (1.1.18).

We have

$$\operatorname{Cov}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix} \operatorname{Cov}\big[\,H_R(\boldsymbol{x}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,H_R(\boldsymbol{x}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big] \\ & \\ \operatorname{Cov}\big[\,H_R(\boldsymbol{y}),\nabla\Phi^N_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,H_R(\boldsymbol{y}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big] \end{bmatrix}$$

$$= \begin{bmatrix} \operatorname{Cov}\big[\,H_R(0),\nabla\Phi^R_{\mathfrak{a}}(0)\,\big] & \operatorname{Cov}\big[\,H_R(\boldsymbol{x}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big] \\ & \\ \operatorname{Cov}\big[\,H_R(\boldsymbol{y}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & \operatorname{Cov}\big[\,H_R(0),\nabla\Phi^R_{\mathfrak{a}}(0)\,\big] \end{bmatrix}.$$

The covariance $\operatorname{Cov}\big[\,H_R(0),\nabla\Phi^R_{\mathfrak{a}}(0)\,\big]$ involves only third order partial derivatives of $\boldsymbol{K}^N_{\mathfrak{a}}$ at 0, and these are all trivial since $\boldsymbol{K}^R_{\mathfrak{a}}$ is an even function. Hence

$$\operatorname{Cov}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix} 0 & \operatorname{Cov}\big[\,H_R(\boldsymbol{x}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big] \\ & \\ \operatorname{Cov}\big[\,H_R(\boldsymbol{y}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & 0 \end{bmatrix}.$$

Similarly

$$\operatorname{Cov}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] = \begin{bmatrix} \operatorname{Cov}\big[\,H_R(\boldsymbol{x}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{x})\,\big] & 0 \\ & \\ 0 & \operatorname{Cov}\big[\,H_R(\boldsymbol{y}),\nabla\Phi^R_{\mathfrak{a}}(\boldsymbol{y})\,\big] \end{bmatrix} = 0.$$

Lemma 2.5.11(ii) implies that

$$\big\|\operatorname{Cov}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]\big\|_{\mathrm{op}} = O\big(\,T_R(\boldsymbol{x}-\boldsymbol{y})\,\big),$$
$$\big\|\operatorname{Cov}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]\big\|_{\mathrm{op}} = O\big(\,T_R(\boldsymbol{x}-\boldsymbol{y})\,\big).$$

Since $\operatorname{Var}\big[\,\nabla\widetilde{\Phi}^N(\boldsymbol{x},\boldsymbol{y})\,\big]$ and we deduce from the regression formula (1.1.18) that

$$\operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big] = \operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big] + O\big(\,T_R(\boldsymbol{x}-\boldsymbol{y})\,\big),$$
$$\operatorname{Var}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big] = \operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big] + O\big(\,T_R(\boldsymbol{x}-\boldsymbol{y})\,\big).$$

The regression formula (1.1.18) shows that

$$\operatorname{Var}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big] = \operatorname{Var}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big]$$
$$-\operatorname{Cov}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\operatorname{Cov}\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y}),\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big].$$
$$= \operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big] + O\big(\,T_R(\boldsymbol{x}-\boldsymbol{y})\,\big)$$
$$-\operatorname{Cov}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\operatorname{Cov}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y}),\widehat{H}^R(\boldsymbol{x},\boldsymbol{y})\,\big].$$

Since $\operatorname{Cov}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big] = O\big(\,T_R(\boldsymbol{x}-\boldsymbol{y})\,\big)$ we deduce from (2.5.37) and (2.5.38) that there exists $C_1>0$, independent of $R>R_0$, such that

$$\operatorname{Cov}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y}),\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\operatorname{Cov}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y}),\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big]$$
$$= O\big(\,T_R(\boldsymbol{x},\boldsymbol{y})\,\big),\;\;\forall \boldsymbol{x},\boldsymbol{y}\in RB,\;\;|\boldsymbol{x}-\boldsymbol{y}|_\infty > C_1,$$

and thus

$$\begin{aligned}&\big\| \operatorname{Var}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big] - \operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big]\,\big\|_{\mathrm{op}}\\ &= O\big(T_R(\boldsymbol{x}-\boldsymbol{y})\big),\ \ \forall \boldsymbol{x},\boldsymbol{y}\in RB,\ \ |\boldsymbol{x}-\boldsymbol{y}|_\infty > C_2 = \max(C_0,C_1).\end{aligned}$$

Since $\operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})\,\big] = \operatorname{Var}\big[\,H_R(0)\,\big]\oplus\operatorname{Var}\big[\,H_R(0)\,\big]$ we deduce from Lemma 2.5.11(i) that there exists $\mu_0>0$ such that

$$\operatorname{Var}\big[\,\widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big] \geq \mu_0\mathbb{1},\ \ \forall R\geq R_0.$$

Note also that (2.5.35) implies that there exists $C_3>0$, independent of $R>R_0$, such that

$$\sup_{\substack{\boldsymbol{x},\boldsymbol{y}\in RB\\ |\boldsymbol{x}-\boldsymbol{y}|_\infty > C_3}} \|\operatorname{Var}\big[\,\widehat{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat\,\big]\|_{\mathrm{op}} = O(1)$$

Lemma 1.1.27 implies that

$$\Big|\,\mathbb{E}\big[\,|\det \widehat{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat|\,\big] - \mathbb{E}\big[\,|\det \widetilde{H}_R(\boldsymbol{x},\boldsymbol{y})^\flat|\,\big]\,\Big| = O\big(T_R(\boldsymbol{x}-\boldsymbol{y})^{1/2}\big). \tag{2.5.40}$$

Using (2.5.38) we deduce that there exists $C_4>0$, independent of $R>R_0$, such that

$$\begin{aligned}&\Big|\,p_{\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})}(0) - p_{\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})}(0)\,\Big|\\ &= \frac{1}{(2\pi)^{m/2}}\Big|\,\det\operatorname{Var}\big[\,\nabla\widehat{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1} - \det\operatorname{Var}\big[\,\nabla\widetilde{\Phi}^R(\boldsymbol{x},\boldsymbol{y})\,\big]^{-1}\,\Big|\\ &= O\big(T_R(\boldsymbol{x}-\boldsymbol{y})\big),\ \ \forall \boldsymbol{x},\boldsymbol{y}\in RB,\ \ |\boldsymbol{x}-\boldsymbol{y}|_\infty > C_4.\end{aligned} \tag{2.5.41}$$

We can now estimate the right-hand-side of (2.5.33). For any $\boldsymbol{x},\boldsymbol{y}\in RB$

$$O\big(T_R(\boldsymbol{x}-\boldsymbol{y})\big) \stackrel{(2.5.34)}{=} O\big(\,|\,\boldsymbol{x}-\boldsymbol{y}\,|_\infty^{-p/2}\big),\ \ \forall p>0.$$

Using (2.5.38), (2.5.39), (2.5.40) and (2.5.41) that we conclude that there exists $C_5>1$, independent of $R>R_0$ such that, for any $p>m$,

$$\forall \boldsymbol{x},\boldsymbol{y}\in RB,\ \ |\boldsymbol{x}-\boldsymbol{y}|_\infty > C_5,\ \ \big|\,\underbrace{\widehat{\rho}_R(\boldsymbol{x},\boldsymbol{y}) - \widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})}_{=\Delta_R(\boldsymbol{x},\boldsymbol{y})}\,\big| = O\big(\,|\,\boldsymbol{x}-\boldsymbol{y}\,|_\infty^{-p/2}\big). \tag{2.5.42}$$

Since the random function $\Phi^R_{\mathfrak{a}}$ is stationary, we deduce that for any $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathbb{R}^m$ such that $\boldsymbol{x}\neq\boldsymbol{y}$ we have

$$\Delta_R(\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}+\boldsymbol{z}) = \Delta_R(\boldsymbol{x},\boldsymbol{y})$$

so $\widehat{\rho}_R(\boldsymbol{x},\boldsymbol{y})$, $\widetilde{\rho}_R(\boldsymbol{x},\boldsymbol{y})$ and $\Delta_R(\boldsymbol{x},\boldsymbol{y})$ depend only on $\boldsymbol{y}-\boldsymbol{x}$.

4. Conclusion Assume now that $\boldsymbol{x},\boldsymbol{y}\in RB$ and $|\boldsymbol{x}-\boldsymbol{y}|_\infty\leq C_5$. Denote by $\widehat{\mathfrak{X}}$ the radial-blowup of $\mathbb{R}^m\times\mathbb{R}^m$ along the diagonal. It is diffeomorphic to the product $\mathbb{R}^m\times S^{m-1}\times[0,\infty)$.

Choose new orthogonal coordinates (ξ,η) given by

$$\xi = \boldsymbol{x}+\boldsymbol{y},\ \ \eta = \boldsymbol{x}-\boldsymbol{y} \Longleftrightarrow \boldsymbol{x} = \frac{1}{2}(\xi+\eta),\ \ \boldsymbol{y} = \frac{1}{2}(\xi-\eta)$$

then

$$|x-y| = |\eta|,\ \ d\boldsymbol{x}d\boldsymbol{y} = 2^{-2m}d\xi d\eta.$$

Note that if $\boldsymbol{x},\boldsymbol{y}\in\operatorname{supp} f_R$, then $|\boldsymbol{x}|,|\boldsymbol{y}|<Rr_0/2$ and thus

$$\boldsymbol{x},\boldsymbol{y}\in\operatorname{supp} f_R \Rightarrow |\xi|,\ |\eta| < \frac{1}{2}|\xi+\eta| + \frac{1}{2}|\xi-\eta| \leq Rr_0. \tag{2.5.43}$$

The natural projection $\pi : \widehat{\mathfrak{X}} \to \mathbb{R}^m \times \mathbb{R}^m$ can given the explicit description

$$\mathbb{R}^m \times S^{m-1} \times [0,\infty) \ni (\xi, \boldsymbol{\nu}, r) \mapsto (\xi, \eta) = (\xi, r\boldsymbol{\nu}) \in \mathbb{R}^m \times \mathbb{R}^m.$$

Set for $R \in (R_0, \infty]$ we set

$$w_R(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|^{m-2} \widehat{\rho}_R(\boldsymbol{x}, \boldsymbol{y}).$$

Lemma 2.5.11(i) implies that for any $C > 0$

$$\sup_{R \in (R_0, \infty]} \sup \| K^R_{\mathfrak{a}} \|_{C^6(RB)} < \infty.$$

We deduce from Proposition 2.4.1 and Lemma 2.5.11 that

$$\sup_{R \in (R_0, \infty]} \sup_{\substack{\boldsymbol{x}, \boldsymbol{y} \in RB \\ 0 < |\boldsymbol{x} - \boldsymbol{y}| \leq C_5}} \big| w_R(\boldsymbol{x}, \boldsymbol{y}) \big| < \infty. \tag{2.5.44}$$

It is easy to see that $\widetilde{\rho}_R \circ \pi$ admits a continuous extension to the blow-up. Using (2.5.42) and (2.5.44) we deduce that for any $p > 0$ there exists a constant $K_p > 0$, independent of R, such that

$$|x - y|^{m-1} \big| \Delta_R(\boldsymbol{x}, \boldsymbol{y}) \big| \leq K_p \big(1 + |x - y| \big)^{-p+m-1}, \;\; \forall \boldsymbol{x}, \boldsymbol{y} \in RB \tag{2.5.45}$$

Set

$$\delta_R(\xi, \eta) = \Delta_R\big(\pi(\xi, \eta) \big)$$

Since $\Delta_R(\boldsymbol{x}, \boldsymbol{y})$ depends only on $\boldsymbol{y} - \boldsymbol{x}$ we deduce that $\delta_R(\xi, \eta)$ is independent of ξ. We have

$$I(R) = \int_{\mathfrak{X}} \Delta_R(\boldsymbol{x}, \boldsymbol{y}) f_R^{\boxtimes 2}(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{x} d\boldsymbol{y} = \int_{\substack{|\boldsymbol{x}|, |\boldsymbol{y}| \leq R r_0/2 \\ \boldsymbol{x} \neq \boldsymbol{y}}} \Delta_R(\boldsymbol{x}, \boldsymbol{y}) f_R^{\boxtimes 2}(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{x} d\boldsymbol{y}$$

$$\stackrel{(2.5.43)}{=} \frac{1}{2^{2m}} \int_{\substack{|\xi| < R r_0, \\ |\boldsymbol{\nu}| = 1, \, r \in (0, R r_0)}} r^{m-1} \delta_R\big(\xi, r\boldsymbol{\nu} \big) f_R\Big(\frac{\xi + r\boldsymbol{\nu}}{2} \Big) f_R\Big(\frac{\xi - r\boldsymbol{\nu}}{2} \Big) dr \, \mathrm{vol}_{S^{m-1}}[d\boldsymbol{\nu}] d\xi$$

$(\xi = 2R\zeta, \; \delta_R(\xi, r\nu) = \delta_R(0, r\nu))$

$$= \Big(\frac{R}{2} \Big)^m \underbrace{\int_{|\zeta| \leq r_0/2} \left(\int_{\substack{|\boldsymbol{\nu}| = 1 \\ r \in (0, R r_0)}} r^{m-1} \delta_R\big(0, r\boldsymbol{\nu} \big) f\Big(\zeta + \frac{r\boldsymbol{\nu}}{2R} \Big) f\big(\zeta - \frac{r\boldsymbol{\nu}}{2R} \big) dr \, \mathrm{vol}_{S^{m-1}}[d\boldsymbol{\nu}] \right) d\zeta}_{=: J(R)}.$$

Note that

$$\delta_R\big(0, r\boldsymbol{\nu} \big) = \widehat{\rho}_R(r\boldsymbol{\nu}/2, -r\boldsymbol{\nu}/2) - \widetilde{\rho}_R(r\boldsymbol{\nu}/2, -r\boldsymbol{\nu}/2)$$

and for $r > 0$, $|\boldsymbol{\nu}| = 1$ fixed

$$\lim_{R \to \infty} \delta_R\big(0, r\boldsymbol{\nu} \big) = \delta_\infty(0, r\nu) = \widehat{\rho}_\infty(r\boldsymbol{\nu}/2, -r\boldsymbol{\nu}/2) - \widetilde{\rho}_\infty(r\boldsymbol{\nu}/2, -r\boldsymbol{\nu}/2).$$

We deduce from (2.5.43) and (2.5.44) that for any $p > 0$ there exists $K_p > 0$ such that for any $R > R_0$, $|\zeta| < r_0/2$, $|\boldsymbol{\nu}| = 1$ and $r \leq R r_0$ we have

$$\Big| r^{m-1} \delta_R\big(0, r\boldsymbol{\nu} \big) f\Big(\zeta + \frac{r\boldsymbol{\nu}}{2R} \Big) f\Big(\zeta - \frac{r\boldsymbol{\nu}}{2R} \Big) \Big| \leq K_p \|f\|^2 \big(1 + r \big)^{-p+m-1}.$$

The constraint $r < R r_0$ is not really necessary since, according to (2.5.43) the left-hand side of the above inequality vanishes if $r > R_0$, $|\zeta| < r_0/2$ and $|\boldsymbol{\nu}| = 1$. For $p > m$ we have

$$\int_{|\zeta| \leq r_0/2} \left(\int_{(0, \infty \times S^{m-1}} \big(1 + r \big)^{-p+m-1} dr \, \mathrm{vol}_{S^{m-1}} \big[d\boldsymbol{\nu} \big] \right) d\zeta < \infty.$$

The dominated convergence theorem implies that $J(R)$ has a finite limit as $R \to \infty$. More precisely

$$\lim_{R\to\infty} J(R) = \int_{|\zeta|\leq r_0/2} \left(\int_{\substack{|\boldsymbol{\nu}|=1 \\ r>0}} r^{m-1}\delta_\infty\big(0, r\boldsymbol{\nu}\big) f(\zeta)^2 dr \operatorname{vol}_{S^{m-1}}[d\boldsymbol{\nu}] \right) d\zeta.$$

This concludes the proof of Theorem 2.5.9. □

Chapter 3

Central limit theorems

3.1. Gaussian Hilbert spaces

3.1.1. Basic definitions and examples.

Definition 3.1.1. A *Gaussian linear space* is a real vector space $\mathfrak{X}$ consisting of (real) Gaussian random variables defined on the same probability space $(\Omega, \mathcal{S}, \mathbb{P})$. If the vector space $\mathfrak{X}$ is closed in $L^2(\Omega, \mathcal{S}, \mathbb{P})$, then we say that $\mathfrak{X}$ is a *Gaussian Hilbert space.* □

Example 3.1.2. Suppose that $\boldsymbol{X}$ is a Fréchet space with dual $\boldsymbol{X}^*$ and $\boldsymbol{\Gamma}$ is a Gaussian measure on $\boldsymbol{X}$. Then the map

$$\mathbf{Ev} : \boldsymbol{X} \times \boldsymbol{X}^* \to \mathbb{R}, \;\; \mathbf{Ev}_x(\xi) = \xi(x)$$

is a centered Gaussian process parametrized by $\boldsymbol{X}^*$. The associated Gaussian Hilbert space $\boldsymbol{X}^*_{\boldsymbol{\Gamma}}$ is the closure in $L^2(\boldsymbol{X}, \boldsymbol{\Gamma})$ of the range the tautological map $T_{\boldsymbol{\Gamma}} : \boldsymbol{X}^* \to L^2\big(\boldsymbol{X}, \boldsymbol{\Gamma}\big)$ defined in (1.1.28). □

Example 3.1.3 (The Main Example)**.** Suppose that

$$X : \big(\Omega, \mathcal{S}, \mathbb{P}\big) \times T \to \mathbb{R}, \;\; (\omega, t) \mapsto X_\omega(t)$$

is a centered Gaussian field parameterized by the set T. The closure in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ of the vector spaces spanned by the collection $\big(X(t)\big)_{t\in T}$ is called the *Gaussian Hilbert space associated to the centered Gaussian field* $(X_t)_{t\in T}$.

□

Definition 3.1.4. An *isonormal Gaussian process* is a triplet $(H, \mathfrak{X}, W)$, where $\mathfrak{X}$ is a Gaussian Hilbert space, H is a Hilbert space and $W : H \to \mathfrak{X}$ is an isomorphism of Hilbert spaces. The map W is called the *white noise map* of the isonormal process. □

Remark 3.1.5. If $(H, \mathfrak{X}, W)$ is an isonormal Gaussian process, then the collection $\big(W(h)\big)_{h\in H}$ is a centered Gaussian process parametrized by H. Its covariance kernel is

$$K : H \times H \to \mathbb{R}, \;\; K(h_1, h_2) = \big(h_1, h_2\big)_H, \;\; \forall h_1, h_2 \in H.$$

Conversely, given a Hilbert space we can use Kolmogorv's existence theorem to conclude that there exist Gaussian processes W on H with the above covariance kernel. We denote by $\mathfrak{X}$ the Gaussian Hilbert space associated to this process as in Example 3.1.3 . Then the induced map $W : H \to \mathfrak{X}$ is linear. Indeed, for any $h, h' \in H$

$$\mathbb{E}\big[\,\big(W(h+h') - W(h) - W(h')\,\big)\big(W(h+h') - W(h) - W(h')\,\big)\,\big] = 0.$$

This proves that the resulting triplet $(H, W, \mathfrak{X})$ is an isonormal Gaussian process. □

Example 3.1.6. To a separable Gaussian space $\mathfrak{X}$ we can non-canonically associate many isonormal processes. Note first that there is a tautological one $(\mathfrak{X}, \mathfrak{X}, \mathbb{1})$.

A complete orthonormal basis $(X_n)_{n\in\mathbb{N}}$ determines a Hilbert space isomorphism

$$\text{Fourier} : \ell_2 \to \mathfrak{X}, \;\; \ell_2 \ni \underline{c} = (c_n)_{n\in\mathbb{N}} \mapsto \text{Fourier}\big(\,\underline{c}\,\big) := \sum_{n\in\mathbb{N}} c_n X_n.$$

□

Example 3.1.7. Let us analyze a special case of the gaussian space described in Example 3.1.2.

Suppose that H is a separable, real Hilbert space with inner product $(-,-)_H$, and Γ is a centered Gaussian measure on H; see Definition 1.1.47.

For any $h \in H$, the linear functional $L_h : H \to \mathbb{R}$, $L_h(x) = (h, x)_H$, is a centered Gaussian random variable. In particular, the collection $(L_h)_{h\in H}$ is a Gaussian random field parameterized by H. We denote by $C(h_1, h_2)$ the covariance of L_{h_1}, L_{h_2},

$$C(h_1, h_2) = \mathbb{E}\big[\,L_{h_1} L_{h_2}\,\big].$$

This defines an inner product on H^*, the topological dual of H. As explained in [**39**], there exists a symmetric, nonnegative *trace class* operator $Q : H \to H$ such that

$$C(h_1, h_2) = \big(\,Qh_1, h_2\,\big)_H, \;\; \forall h_1, h_2 \in H.$$

Assume for simplicity that $\ker Q = 0$.

To this Gaussian measure we can associate the Gaussian Hilbert space $H^*_{\mathbf{\Gamma}}$ defined as the closure in $L^2(H, \mathbf{\Gamma})$ of the vector space spanned by $(L_h)_{h\in H}$. One could think of the elements of $H^*_{\mathbf{\Gamma}}$ as measurable linear functionals $H \to \mathbb{R}$.

Note that we have a continuous map with dense image

$$L : H \to H^*_{\mathbf{\Gamma}}, \;\; h \mapsto L_h. \tag{3.1.1}$$

The Hilbert space $H^*_{\mathbf{\Gamma}}$ is canonically isomorphic with H as a Hilbert space. To construct this isomorphism consider the dense subspace $Q^{1/2}H$ and the map

$$W : Q^{1/2}H \to L^2(H, \mathbf{\Gamma}), \;\; Q^{1/2}H \ni z \mapsto W_z := L_{Q^{-1/2}z}.$$

Clearly the image of W is equal to the image of the map L in (3.1.1). Observe that

$$\mathbb{E}\big[\,W_{z_1} W_{z_2}\,\big] = (z_1, z_2)_H, \;\; \forall z_1, z_2 \in Q^{-1/2}H.$$

This shows that the map W extends by continuity to an isometry $W : H \to H^*_{\mathbf{\Gamma}}$. This isomorphism of Hilbert spaces is called the *white noise map*. Observe that the triplet $(H, H^*_{\mathbf{\Gamma}}, W)$ is an isonormal Gaussian process.

The subspace $Q^{1/2}H \subset H$ is the *Cameron-Martin space* defined in (1.1.34). If we identify H with its topological dual we observe that $H^* = H \subset H^*_{\mathbf{\Gamma}}$ then $Q^{1/2}H$ can be identified with the Cameron-Martin space of the Gaussian process (L_h); see Appendix B.5.

We fix an orthonormal $(e_k)_{k\in\mathbb{N}}$ (complete) basis of H consisting of eigenvectors of Q,

$$Qe_n = \lambda_n e_n, \;\; n \in \mathbb{N}.$$

The collection of linear functionals

$$W_{e_n} := \frac{1}{\sqrt{\lambda_n}} L_{e_n}, \;\; n \in \mathbb{N}$$

is an orthonormal basis of the associated Gaussian Hilbert space $H^*_{\mathbf{\Gamma}}$. □

Definition 3.1.8. Suppose that $\mathfrak{X} \subset L^2(\Omega, \mathcal{S}, \mathbb{P})$ is a Gaussian Hilbert space. We denote by $\mathcal{S}_{\mathfrak{X}}$ the σ-subalgebra of $\mathcal{S}$ generated by the collection of random variables $X \in \mathfrak{X}$ and we define

$$\mathcal{F}(\mathfrak{X}) := L^2\big(\Omega, \mathcal{S}_{\mathfrak{X}}, \mathbb{P}\big) \subset L^2(\Omega, \mathcal{S}, \mathbb{P}).$$

For reasons that will become clear a bit later, we will refer to $\mathcal{F}(\mathfrak{X})$ as the *Wiener chaos* of $\mathfrak{X}$. □

Example 3.1.9. Suppose that T is a compact metric space and $\mathbf{\Gamma}$ is a Gaussian measure on the Banach space $\boldsymbol{X} = C(T)$. We have a Gaussian stochastic process

$$\widehat{X} : \big(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}}, \mathbf{\Gamma}\big) \times T \to \mathbb{R}, \;\; \boldsymbol{X} \times T \ni (f,t) \mapsto \widehat{X}_f(t) := \mathbf{Ev}_t(f) = f(t).$$

The Gaussian Hilbert space $\mathfrak{X}$ determined by this process is the closure in $L^2\big(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}}, \mathbf{\Gamma}\big)$ of the subspace spanned by $\widehat{X}(t)$, $t \in T$. Blackwell's theorem implies that the sigma-algebra generated by $\big(\mathbf{Ev}_t\big)_{t\in T}$ coincides with the Borel sigma-algebra $\mathcal{B}_{\boldsymbol{X}}$ and thus

$$\mathcal{F}(\mathfrak{X}) = L^2\big(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}}, \mathbf{\Gamma}\big).$$

□

3.1.2. Hermite decompositions. To understand what happens when we pass from a Gaussian Hilbert space $\mathfrak{X}$ to its Wiener chaos $\mathcal{F}(\mathfrak{X})$ we consider first the simplest possible case, $\dim \mathfrak{X} = 1$.

Example 3.1.10 (Hermite polynomials)**.** Consider the standard Gaussian measure $\mathbb{P} = \mathbf{\Gamma}$ on $\Omega = \mathbb{R}$,

$$\mathbf{\Gamma}\big[\, dx\,\big] = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \boldsymbol{\lambda}\big[\, dx\,\big].$$

As explained in Example 3.1.2, this tautologically defines a one-dimensional Gaussian Hilbert space $\mathfrak{X}_1$ spanned by the identity function $\mathbb{1}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$.

In this case $\mathcal{S} = \mathcal{S}_{\mathfrak{X}_1}$ is the σ-algebra $\mathcal{B}_{\mathbb{R}}$ of the Borel subsets of $\mathbb{R}$ and $\mathcal{S}_{\mathfrak{X}_1} = \mathcal{B}_{\mathbb{R}}$. Moreover, we have an isomorphism

$$L^2(\mathbb{R}, \mathbf{\Gamma}) \to \mathcal{F}(\mathfrak{X}_1), \;\; L^2(\mathbb{R}, \mathbf{\Gamma}) \ni F \mapsto F \circ \mathbb{1}_{\mathbb{R}}.$$

We see that the Wiener chaos $\mathcal{F}(\mathfrak{X}_1)$ is much larger than $\mathfrak{X}_1$.

A convenient complete orthogonal basis of $\mathcal{F}(\mathfrak{X}_1) = L^2(\mathbb{R}, \mathbf{\Gamma})$ is given by the *Hermite polynomials* $(H_n)_{n\geq 0}$, [**94**, V.1.3].

To define these polynomials we introduce the *creation operator* $\delta_x : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$,

$$\delta_x f(x) = -e^{\frac{x^2}{2}} \partial_x \big(e^{-\frac{x^2}{2}} f(x) \big) = -\partial_x f(x) + x f(x). \tag{3.1.2}$$

The creation operator is the formal adjoint with respect to the inner product in $L^2(\mathbb{R}, \boldsymbol{\Gamma})$ of the *annihilation operator*, namely the usual differential operator ∂_x. More precisely,

$$\int_{\mathbb{R}} f'(x) g(x) \boldsymbol{\Gamma}\big[\, dx \,\big] = \int_{\mathbb{R}} f(x) \delta_x g(x) \boldsymbol{\Gamma}\big[\, dx \,\big], \;\; \forall f, g \in C^\infty_{\mathrm{cpt}}(\mathbb{R}).$$

The above equality is a direct consequence of the Gaussian integration-by-parts formula (1.1.5).

The *n-th Hermite polynomial* is defined by

$$H_n(x) = \delta_x^n 1. \tag{3.1.3}$$

Let us observe that the operators ∂_x, δ_x satisfy the *Heisenberg identity*

$$[\partial_x, \delta_x] = \partial_x \delta_x - \delta_x \partial_x = \mathbb{1}.$$

Using this iteratively we deduce

$$\partial_x H_n(x) = n H_{n-1}(x), \;\; \forall n \in \mathbb{N}, \tag{3.1.4a}$$

$$\delta_x \partial_x H_n(x) = -H_n''(x) + x H_n'(x) = n H_n(x), \;\; \forall n \in \mathbb{N}. \tag{3.1.4b}$$

$$\partial_x^n \big(e^{-\frac{x^2}{2}} \big) = (-1)^n H_n(x) e^{-\frac{x^2}{2}}. \tag{3.1.4c}$$

From the defining equation (3.1.3) we deduce

$$H_n(x) = \delta_x H_{n-1}(x) = -H_{n-1}'(x) + x H_{n-1}(x) \overset{(3.1.4b)}{=} -(n-1) H_{n-2}(x) + x H_{n-1}(x),$$

we thus we obtain the three-term recurrence relations

$$H_{n+1}(x) = \delta_x H_n(x) = -H_n'(x) + x H_n(x), \;\; \forall n \geq 0. \tag{3.1.5a}$$

$$H_n(x) = x H_{n-1}(x) - (n-1) H_{n-2}(x), \;\; \forall n \geq 2. \tag{3.1.5b}$$

For example,

$$H_0(x) = 1, \;\; H_1(x) = x, \;\; H_2(x) = x^2 - 1, \;\; H_3(x) = x^3 - 3x,$$

$$H_4(x) = x^4 - 6x^2 + 3, \;\; H_5(x) = x^5 - 10\, x^3 + 15\, x, \;\; H_6(x) = x^6 - 15\, x^4 + 45\, x^2 - 15.$$

Using the equalities $H_n' = n H_{n-1}$ and $H_n(0) = -(n-1) H_{n-2}(0)$ we deduce inductively that

$$H_n(x) = n! \sum_{r=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^r}{2^r r! (n-2r)!} x^{n-2r}. \tag{3.1.6}$$

Observe that the leading coefficient of $H_n(x)$ is 1. We have

$$|H_n(x)| \leq \sum_{m=0}^{n} \binom{n}{m} |x|^{n-m} \left(\frac{1}{2} \right)^{m/2} = \left(\frac{1}{\sqrt{2}} + |x| \right)^n, \;\; \forall \boldsymbol{x} \in \mathbb{R}. \tag{3.1.7}$$

From the equalities (3.1.3) and (3.1.4a) we deduce that the collection $(H_n)_{n \geq 0}$ is orthogonal in $L^2(\mathbb{R}, \boldsymbol{\Gamma})$,

$$\int_{\mathbb{R}} H_m(x) H_n(x) \boldsymbol{\Gamma}\big[\, dx \,\big] = \delta_{nm} n!. \tag{3.1.8}$$

In particular,

$$\int_{\mathbb{R}} H_m(x)\boldsymbol{\Gamma}\big[\,dx\,\big] = \int_{\mathbb{R}} H_m(x)H_0(x)\boldsymbol{\Gamma}\big[\,dx\,\big] = 0,\ \ \forall m > 0. \tag{3.1.9}$$

Proposition 1.1.3 shows that the collection $(H_n)_{n\geq 0}$ spans a dense subspace in $L^2(\mathbb{R}, \boldsymbol{\Gamma})$. Hence, any $f \in L^2(\mathbb{R}, \boldsymbol{\Gamma})$ admits a *Fourier-Hermite decomposition*

$$f = \sum_{n\geq 0} c_n(f)H_n(x),\ \ c_n(f) = \frac{1}{n!}\int_{\mathbb{R}} f(x)H_n(x)\boldsymbol{\Gamma}\big[\,dx\,\big].$$

Let us point out that if $g \in C^\infty(\mathbb{R})$ has the property that

$$g^{(k)} \in L^2(\mathbb{R}, \boldsymbol{\Gamma}),\ \ \forall k \geq 0,$$

then we have the following expansion in $L^2(\mathbb{R}, \boldsymbol{\Gamma})$

$$g(x) = \sum_{n\geq 0} \frac{1}{n!}\mathbb{E}_{\boldsymbol{\Gamma}}\big[\,g^{(n)}\,\big]H_n(x), \tag{3.1.10}$$

where $\mathbb{E}_{\gamma_1}$ denotes the expectation with respect to the probability measure $\boldsymbol{\Gamma}$. If in the above equality we choose

$$g(x) = g_t(x) = e^{tx-\frac{t^2}{2}},$$

then, for any $t \in \mathbb{R}$, we have

$$\frac{d^n}{dx^n}g_t(x) = t^n e^{tx-\frac{t^2}{2}},\ \ \mathbb{E}_{\boldsymbol{\Gamma}}\big[\,g_t^{(n)}\,\big] = t^n e^{-\frac{t^2}{2}}\int_{\mathbb{R}} e^{tx}\boldsymbol{\Gamma}\big[\,dx\,\big] \overset{(1.1.7)}{=} t^n.$$

This proves that

$$\sum_{n\geq 0} H_n(x)\frac{t^n}{n!} = e^{tx-\frac{t^2}{2}} = g_t(x), \tag{3.1.11}$$

where the above series converges in $L^2(\mathbb{R}, \boldsymbol{\Gamma})$ for any $t \in \mathbb{C}$. The estimates (3.1.7) show that the above series also converges uniformly for (x, t) on the compacts of $\mathbb{R} \times \mathbb{C}$. □

Remark 3.1.11. There is no consensus in the existing literature on the canonical definition of Hermite polynomials since many authors use different normalizations as canonical. To help the reader navigate these "canonical" choices we want to describe a one-parameter family of "canonical" Hermite polynomials that contains most these choices. Our presentation follows closely [**84**, Sec. 9.3] to which we refer for proofs and more details.

For each $\rho > 0$ and $x \in \mathbb{R}$ we set

$$H_n(x|\rho) := (-\rho)^n \frac{d^n}{dx^n} e^{-\frac{x^2}{2\rho}}.$$

The function $H_n(x|\rho)$ is a degree n polynomial in x called *the nth Hermite polynomial with parameter ρ.*

The exponential generating function of the sequence $\big(\,H_n(x, \rho)\,\big)_{n\geq 0}$ is

$$H_\rho(t, x) := e^{tx-\frac{1}{2}\rho t^2} = \sum_{n\geq 0} \frac{1}{n!}H_n(x|\rho)t^n. \tag{3.1.12}$$

Thus $H_n(x) = H_n(x|\rho = 1)$. Moreover,

$$H_n(x|\rho) = \rho^{n/2}H_n\big(\,x\rho^{-\frac{1}{2}}\,\big). \tag{3.1.13}$$

In particular this shows that the leading coefficient of $H_n(x|\rho)$ is 1 for any ρ. Using (3.1.4a) we deduce

$$\partial_x H_n(x|\rho) = nH_{n-1}(x|\rho). \tag{3.1.14}$$

Note that

$$\partial_\rho H_\rho(t,x) = -\frac{t^2}{2}H_\rho(t,x) = -\frac{1}{2}\partial_x^2 H_\rho(t,x).$$

Thus the polynomials $H_n(x|\rho)$ satisfy the backwards heat equation

$$\Big(\partial_\rho + \frac{1}{2}\partial_x^2\Big)H_n(x|\rho) = 0, \;\; \forall n \geq 0. \tag{3.1.15}$$

Using (3.1.12), (3.1.15) and Itô's formula one can show that if $B(t)$ is a one-dimensional Brownian motion started at 0, then, for any $t > 0$,

$$H_{n+1}\big(B(t)\,\big|t\big) = \int_0^t H_n\big(B(s)\,\big|s\big)dB(s). \tag{3.1.16}$$

For details we refer to [**97**, Sec 2.7]. □

Suppose that $\mathfrak{X} \subset L^2(\Omega, \mathcal{S}, \mathbb{P})$ is a *separable* Gaussian Hilbert space. Fix a complete orthonormal base $(X_n)_{n\geq 1}$ of $\mathfrak{X}$. In particular, we have

$$\mathbb{E}\big[\,X_i X_j\,\big] = \delta_{ij} = \text{the Kronecker } \delta,$$

and thus the random variables $(X_n)_{n\geq 1}$ are independent. Additionally, the σ-algebra generated by the collection (X_n) coincides with the σ-algebra $\mathcal{S}_{\mathfrak{X}}$.

Consider the space $\mathbb{R}^{\mathbb{N}}$ of real sequences $\underline{x} = (x_1, x_2, \dots)$ equipped with the product measure

$$\boldsymbol{\Gamma}^{\mathbb{N}} = \bigotimes_{n\in\mathbb{N}} \boldsymbol{\Gamma}\big[\,dx_n\,\big],$$

defined on the Borel σ-algebra $\mathcal{B}^{\mathbb{N}}$ of the space $\mathbb{R}^{\mathbb{N}}$ equipped with the product topology.

For $n \in \mathbb{N}$ we denote by π_n the natural projection

$$\mathbb{R}^{\mathbb{N}} \to \mathbb{R}^n, \;\; \underline{x} \mapsto (x_1, \dots, x_n)$$

and we set $\mathcal{B}_n := \pi_n^{-1}\big(\mathcal{B}_{\mathbb{R}^n}\big)$. Then

$$\mathcal{B}^{\mathbb{N}} = \bigvee_{n\in\mathbb{N}} \mathcal{B}_n.$$

The L^2-martingale convergence theorem implies[1] that the union of the subspaces

$$L^2\big(\mathbb{R}^{\mathbb{N}}, \mathcal{B}_n, \boldsymbol{\Gamma}^{\mathbb{N}}\big), \;\; n \in \mathbb{N},$$

is a dense subspace of $L^2\big(\mathbb{R}^N, \mathcal{B}^{\mathbb{N}}, \boldsymbol{\Gamma}^{\mathbb{N}}\big)$.

We have a natural map

$$\vec{X} : \Omega \to \mathbb{R}^{\mathbb{N}}, \;\; \omega \mapsto \big(\,X_1(\omega), X_2(\omega), \dots,\,\big).$$

Then

$$\mathcal{S}_{\mathfrak{X}} = \sigma(X_1, X_2, \dots,) = \vec{X}^{-1}(\mathcal{B}^{\mathbb{N}}),$$

[1] One could use the Monotone Class Theorem to reach the same conclusion, but the details would fill-up more space.

and $\vec{X}_{\#}(\mathbb{P}) = \boldsymbol{\Gamma}^{\mathbb{N}}$. Moreover (see [**35**, Cor.II.4.5]), a function $f : \Omega \to \mathbb{R}$ is $\mathcal{S}_{\mathfrak{X}}$-measurable if and only if there exists a $\mathcal{B}^{\mathbb{N}}$-measurable function $F : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, such that

$$f(\omega) = F\big(\, X_1(\omega), X_2(\omega), \dots \big), \;\; \forall \omega \in \Omega.$$

Additionally, $f \in \mathcal{F}(\mathfrak{X})$ iff $F \in L^2(\mathbb{R}^{\mathbb{N}}, \boldsymbol{\Gamma}^{\mathbb{N}})$ and

$$\int_{\Omega} f(\omega)^2 \mathbb{P}\big[\, d\omega \,\big] = \int_{\mathbb{R}^{\mathbb{N}}} F(\underline{x})^2 \boldsymbol{\Gamma}^{\mathbb{N}}\big[\, d\underline{x} \,\big].$$

This yields an isomorphism of Hilbert spaces

$$\mathcal{F}(\mathfrak{X}) = L^2\big(\, \Omega, \mathcal{S}_{\mathfrak{X}}, \mathbb{P} \,\big) \to L^2\big(\, \mathbb{R}^{\mathbb{N}}, \boldsymbol{\Gamma}^{\mathbb{N}} \,\big).$$

We can construct an orthonormal basis of $L^2(\mathbb{R}^{\mathbb{N}}, \boldsymbol{\Gamma}^{\mathbb{N}})$ as follows. For any multi-index

$$\alpha = (\alpha_1, \alpha_2, \dots) \in \mathbb{N}_0^{\mathbb{N}}$$

such the $\alpha_k = 0$ for all k sufficiently large, we consider the multi-variable polynomial

$$H_\alpha(\underline{x}) := \prod_{k \in \mathbb{N}} H_{\alpha_k}(x_k), \;\; \underline{x} = (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}}.$$

The collection H_α thus obtained is a complete orthogonal basis of $L^2(\mathbb{R}^{\mathbb{N}}, \boldsymbol{\Gamma}^{\mathbb{N}})$ and (3.1.8) shows that

$$\| \, H_\alpha \, \|^2_{L^2(\mathbb{R}^N, \boldsymbol{\Gamma}^{\mathbb{N}})} = \alpha! := \prod_{k=1}^{\infty} \alpha_k!. \tag{3.1.17}$$

3.1.3. Wick's formula. Suppose that $X_1, \dots, X_n$ are jointly Gaussian, centered real random variables, i.e., the real random vector $\boldsymbol{X} = (X_1, \dots, X_n)$ is centered Gaussian. Wick's formula provides an explicit description of the expectation $\mathbb{E}\big[\, X_1 \cdots X_n \,\big]$ in terms of the variance operator $\operatorname{Var}\big[\, \boldsymbol{X} \,\big]$ of the Gaussian vector $\boldsymbol{X}$.

Let us first observe that

$$(-1)^n \mathbb{E}\big[\, X_1 \cdots X_n \,\big] = \mathbb{E}\big[\, (-X_1) \cdots (-X_n) \,\big] = \mathbb{E}\big[\, X_1 \cdots X_n \,\big],$$

where the second equality is due to the symmetry of distribution $\mathbb{P}_{\boldsymbol{X}}$ of $\boldsymbol{X}$, i.e., for any Borel subset B of $\mathbb{R}^n$

$$\mathbb{P}\big[\, \{\boldsymbol{X} \in -B\} \,\big] = \mathbb{P}\big[\, \{\boldsymbol{X} \in B\} \,\big].$$

In particular, this shows that

$$\mathbb{E}\big[\, X_1 \cdots X_n \,\big] = 0 \;\text{ if } n \text{ is odd}. \tag{3.1.18}$$

To explain how to compute the expectation $\mathbb{E}\big[\, X_1 \cdots X_n \,\big]$ in terms of the covariances $\mathbb{E}\big[\, X_i X_j \,\big]$ we need a bit of combinatorial terminology.

Let V be a finite set. A *Feynman diagram* on V is a graph with vertex set V such that any vertex is connected to at most one other vertex. In other words, a Feynman diagram is a partial matching of the vertices in V. We denote $\operatorname{Feyn}(V)$ the set of Feynman diagrams with vertex set V.

Given $\Gamma \in \operatorname{Feyn}(V)$ we denote by $\mathcal{E}(\Gamma)$ the set of edges of Γ and by $\mathcal{I}(\Gamma)$ the set of isolated vertices Γ, i.e., vertices not connected to any other vertex. The *rank* of a Feynman diagram is the number of its edges, $r(\Gamma) := \#\mathcal{E}(\Gamma)$. We have

$$\#V = 2r(\Gamma) + \#\mathcal{I}(\Gamma).$$

A diagram is called *complete* if it has no isolated vertices, $\mathcal{I}(\Gamma) = \emptyset$.

We denote by $\operatorname{Feyn}^r(V)$ the subset of $\operatorname{Feyn}(V)$ consisting of diagrams of rank r. We denote by $\operatorname{Feyn}^*(V)$ the set of complete Feynman diagrams. We set $\mathbb{I}_n := \{1, \dots, n\}$ and

$$\operatorname{Feyn}(n) := \operatorname{Feyn}(\mathbb{I}_n), \ \ \operatorname{Feyn}^r(n) := \operatorname{Feyn}^r(\mathbb{I}_n) \text{ etc.}$$

Lemma 3.1.12.

$$\#\operatorname{Feyn}^*(n) = \begin{cases} 0, & n = 2m+1, \\ (2m-1)!!, & n = 2m \end{cases}$$

$$\#\operatorname{Feyn}^r(n) = \binom{n}{n-2r} \times \#\operatorname{Feyn}^*(2r) = \binom{n}{n-2r}(2r-1)!! = \frac{n!}{2^r r!(n-2r)!}.$$

Proof. Only the case $n = 2m$ is nontrivial. Here is how one generates all the complete diagrams with $2m$ vertices $X_1, \dots, X_{2m}$.

Take the vertex X_1 and pair it with one of the remaining $(2m-1)$ vertices. There are $(2m-1)$ possibilities. Once X_1 is paired, we are left with $(2m-2)$ vertices and there are $\#\operatorname{Feyn}_*(2m-2)$ complete Feynman diagrams on $(2m-2)$ vertices. Hence

$$\#\operatorname{Feyn}^*(2m) = (2m-1) \times \#\operatorname{Feyn}^*(2m-2).$$

In general

$$\#\operatorname{Feyn}^r(n) = \sum_{\substack{S \subset \mathbb{I}_n \\ \#S = n-2r}} \#\operatorname{Feyn}^*\left(\mathbb{I}_n \setminus S\right) = \binom{n}{n-2r} \times \#\operatorname{Feyn}^*(2r).$$

□

If we set

$$d_n(r) := \#\operatorname{Feyn}^r(n),$$

we deduce from (3.1.6) that

$$H_n(x) = \sum_{r=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^r \sum_{\Gamma \in \operatorname{Feyn}^r(n)} x^{n-2r} = \sum_{r=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^r d_n(r) x^{n-2r}. \tag{3.1.19}$$

Suppose we are given a jointly Gaussian family of centered random variables $(X_v)_{v \in V}$. For $\Gamma \in \operatorname{Feyn}(V)$ we define the *random variable*

$$w(\Gamma) = w(\Gamma)\big[\,(X_v)_{v \in V}\,\big] := \left(\prod_{\boldsymbol{e} \in \mathcal{E}(\Gamma)} w(\boldsymbol{e})\right) \cdot \prod_{v \in \mathcal{I}(\Gamma)} X_v,$$

where for any edge $\boldsymbol{e} = [v_1, v_2]$ of Γ we define its weight to be the covariance $w(\boldsymbol{e}) = \mathbb{E}\big[\,X_{v_1} X_{v_2}\,\big]$. Note that if Γ is complete, then $w(\Gamma)$ is deterministic, i.e., a real constant. If $\emptyset$ denotes the Feynman diagram with no edges, then

$$w(\emptyset) = \prod_{v \in V} X_v.$$

Proposition 3.1.13 (Wick's formula)**.** *Suppose that V is a finite set and $(X_v)_{v\in V}$ is a jointly Gaussian family of centered random variables. Then*

$$\mathbb{E}\big[\, w(\emptyset)\,\big] = \sum_{\Gamma\in \mathrm{Feyn}^*(V)} w(\Gamma) \tag{3.1.20}$$

Proof. We can assume that $V = \mathbb{I}_n$, $n = \#V$. Note that if n is odd, then the sum in the right-hand-side of (3.1.20) is trivial. This agrees with (3.1.18)

To prove (3.1.20) we first observe that

$$\mathbb{E}[X_1\cdots X_n] = \frac{1}{n!}\frac{\partial^n}{\partial t_1\cdots\partial t_n}\Big|_{t_1=\cdots=t_n=0}\mathbb{E}\big[\,(t_1X_1+\cdots+t_nX_n)^n\,\big].$$

Next, we observe that $t_1X_1+\cdots+t_nX_n$ is a centered Gaussian variable with variance

$$v(t_1,\dots,t_n) = \sum_{i,j=1}^n \mathbb{E}[X_iX_j]t_it_j = \sum_{j=1}^n \mathbb{E}[X_j^2]t_j^2 + 2\sum_{i<j}\mathbb{E}[X_iX_j]t_it_j.$$

If we let $n = 2k$, we deduce from (1.1.9) that

$$\mathbb{E}\big[\,(t_1X_1+\cdots+t_{2k}X_{2k})^{2k}\,\big] = (2k-1)!!\left(\sum_{i,j=1}^{2k}\mathbb{E}[X_iX_j]t_it_j\right)^k.$$

so that

$$\mathbb{E}[X_1\cdots X_{2k}] = \frac{(2k-1)!!}{(2k)!}\frac{\partial^{2k}}{\partial t_1\cdots\partial t_{2k}}\Big|_{t_1=\cdots=t_{2k}=0}\left(\sum_{i,j=1}^{2k}\mathbb{E}[X_iX_j]t_it_j\right)^k$$

$$= (2^kk!)\frac{(2k-1)!!}{(2k)!}\sum_{\Gamma\in\mathrm{Feyn}^*(n)} w(\Gamma) = \sum_{\Gamma\in\mathrm{Feyn}^*(n)} w(\Gamma).$$

□

Example 3.1.14. Suppose that the random variables Y_1, Y_2 are centered and jointly Gaussian. We set $v_i = \mathbb{E}\big[\,Y_i^2\,\big]$, $c = \mathbb{E}\big[\,Y_1Y_2\,\big]$. Applying Wick's formula to the Gaussian vetor $X = (Y_1, Y_1, Y_1, Y_2)$ we deduce

$$\mathbb{E}\big[\,Y_1^3Y_2\,\big] = 3v_1c.$$

Similarly

$$\mathbb{E}\big[\,Y_1^2Y_2^2\,\big] = v_1v_2 + 2c^2.$$

□

3.1.4. The Wiener chaos decomposition. Fix a probability space $(\Omega, \mathcal{S}, \mathbb{P})$ and a separable Gaussian Hilbert space $\mathfrak{X}\subset L^2(\Omega,\mathcal{S},\mathbb{P})$. We want to describe a coordinate independent orthogonal decomposition of the Wiener chaos $\mathcal{F}(\mathfrak{X})$ that is closely related to the coordinate dependent Hermite decomposition described in Subsection 3.1.2.

Proposition 3.1.15. *The vector space*

$$\mathrm{span}_{\mathbb{R}}\big\{\,\xi_1\cdots\xi_n;\quad n\in\mathbb{N},\ \ \xi_1,\dots,\xi_n\in\mathfrak{X}\,\big\}$$

is dense in $\mathcal{F}(\mathfrak{X}) = L^2(\Omega,\mathcal{S}_{\mathfrak{X}},\mathbb{P})$.

Proof. Fix a complete orthonormal basis $X_1, X_2, \dots, X_n, \dots$ of $\mathfrak{X}$. I will prove that

$$\mathcal{P}_{\mathbb{C}} := \operatorname{span}_{\mathbb{C}}\Big\{ X_1^{\alpha_1}\cdots X_n^{\alpha_n};\quad n\in\mathbb{N},\ \ \alpha_1,\dots,\alpha_n\in\mathbb{N}_0 \Big\}$$

is dense in $\mathcal{F}(\mathfrak{X})_{\mathbb{C}} = \mathcal{F}(\mathfrak{X}) \oplus \boldsymbol{i}\mathcal{F}(\mathfrak{X})$. I follow the approach in [**76**, Thm.2.6].

Denote by $\widehat{\mathcal{P}}_{\mathbb{C}}$ the closure of $\mathcal{P}_{\mathbb{C}}$ in $\mathcal{F}(\mathfrak{X})_{\mathbb{C}}$. We will prove prove that $\mathcal{F}(\mathfrak{X})_{\mathbb{C}} \subset \widehat{\mathcal{P}}_{\mathbb{C}}$. Set

$$\mathcal{V}_n = \operatorname{span}\Big\{ X_1, X_2, \dots, X_n, \Big\},\ \ \mathcal{V} = \bigcup_{n\geq 1}\mathcal{V}_n.$$

The result follows from the following two facts.

A. $e^{\boldsymbol{i}X} \in \mathcal{F}(\mathfrak{X})_{\mathbb{C}} = \mathcal{F}(\mathfrak{X}) + \boldsymbol{i}\mathcal{F}(\mathfrak{X})$ for any $X\in\mathfrak{X}$.

Proof. Let $X\in\mathfrak{X}$. Then

$$e^{\boldsymbol{i}X} = \sum_{k=0}^{\infty}\frac{\boldsymbol{i}^k}{k!}X^k,$$

where the above series converges in $L^2(\Omega,\mathcal{S},\mathbb{P})$. This proves that $e^{\boldsymbol{i}X}\in\mathcal{F}(\mathfrak{X})_{\mathbb{C}}$.

B. If $Z\in\mathcal{F}(\mathfrak{X})_{\mathbb{C}}$ and $\mathbb{E}\big[\, Ze^{\boldsymbol{i}X}\,\big] = 0$, $\forall X\in\mathcal{V}$, then $Z=0$.

Proof. We set

$$\mathcal{F}_n := \sigma(X_1, X_2, \dots, X_n),$$

so we get a filtration of σ-algebras $\mathcal{F}_1\subset\mathcal{F}_2\subset\cdots$ such that

$$\mathcal{S}_{\mathfrak{X}} = \bigvee_{n=1}^{\infty}\mathcal{F}_n. \tag{3.1.21}$$

Suppose that $Z\in\mathcal{F}(\mathfrak{X})_{\mathbb{C}}$ and $\mathbb{E}\big[\, Ze^{\boldsymbol{i}X}\,\big]=0$, $\forall X\in\mathfrak{X}$. We set

$$Z_n := \mathbb{E}\big[\, Z \,\|\, \mathcal{F}_n\,\big].$$

The definition of conditional expectation implies that

$$\mathbb{E}[\, Z_n e^{\boldsymbol{i}X}\,] = 0,\ \ \forall X\in\mathcal{V}_n.$$

Now observe that since $Z_n\in L^2(\Omega,\mathcal{F}_n,\mathbb{P})$ we have

$$Z_n(\omega) = \varphi_n(\, X_1(\omega),\dots,X_n(\omega)\,)$$

for some $\varphi\in L^2(\mathbb{R}^n,\boldsymbol{\Gamma}^n)$. We deduce that

$$\mathbb{E}\big[\,\varphi_n(X_1,\dots,X_n)e^{\boldsymbol{i}t_1X_1+\cdots+\boldsymbol{i}t_nX_n}\,\big] = 0,\ \ \forall t_1,\dots,t_n\in\mathbb{R}.$$

In other words, the Fourier transform of the complex valued measure

$$\varphi_n(x_1,\dots,x_n)\boldsymbol{\Gamma}\big[\,dx_1\,\big]\cdots\boldsymbol{\Gamma}_n\big[\,dx_n\,\big]$$

is trivial so that $\varphi_n=0$. Hence $Z_n=0$, $\forall n\in\mathbb{N}$, i.e.,

$$\mathbb{E}\big[\, Z\,\|\,\mathcal{F}_n\,\big] = 0,\ \ \forall n\in\mathbb{N}.$$

Using (3.1.21), we deduce from the Martingale Convergence Theorem

$$Z = \mathbb{E}\big[\, Z\,\|\,\mathcal{S}_{\mathfrak{X}}\,\big] = \lim_{n\to\infty}\mathbb{E}\big[\, Z\,\|\,\mathcal{F}_n\,\big] = 0.$$

□

For $n \in \mathbb{N}_0$ we define $\mathcal{P}_n(\mathfrak{X})$ to be the closure in $\mathcal{F}(\mathfrak{X})$ of the subspace

$$\left\{ p(\xi_1, \dots, \xi_m);\ \ m > 0,\ \ \xi_1, \dots, \xi_m \in \mathfrak{X},\ \ p \in \mathbb{R}[x_1, \dots, x_m],\ \ \deg p \leq n \right\}.$$

Proposition 3.1.15 shows that the vector space

$$\mathcal{P}(X) = \bigcup_{n \geq 0} \mathcal{P}_n(\mathfrak{X}),$$

is dense in $\mathcal{F}(\mathfrak{X})$. Clearly $\mathcal{P}_{n-1}(\mathfrak{X}) \subset \mathcal{P}_n(\mathfrak{X})$. We denote by $\mathfrak{X}^{:n:}$ the orthogonal complement of $\mathcal{P}_{n-1}(\mathfrak{X})$ in $\mathcal{P}_n(\mathfrak{X})$. We deduce that

$$\mathcal{F}(\mathfrak{X}) = \widehat{\bigoplus_{n \geq 0}} \mathfrak{X}^{:n:}, \tag{3.1.22}$$

where the direct sum in the right-hand-side indicates a Hilbert-complete direct sum, i.e.,

$$\xi \in \widehat{\bigoplus_{n \geq 0}} \mathfrak{X}^{:n:} \Longleftrightarrow \xi = (\xi_n)_{n \geq 0},\ \ \xi_n \in \mathfrak{X}^{:n:},\ \ \sum_{n \geq 0} \|\xi_n\|^2_{L^2} < \infty.$$

The decomposition (3.1.22) is called the *Wiener chaos decomposition* of $\mathcal{F}(\mathfrak{X})$. We will denote by Proj_n the orthogonal projection $\mathcal{F}(\mathfrak{X}) \to \mathfrak{X}^{:n:}$. Note that

$$\mathfrak{X}^{:0:} = \mathrm{span}\{1\}$$

so

$$\forall n > 0,\ \ \forall F \in \mathfrak{X}^{:n:},\ \ \mathbb{E}\big[\, F \,\big] = \mathbb{E}\big[\, F \cdot 1 \,\big] = 0.$$

Example 3.1.16. Suppose that $\mathfrak{X}$ is the 1-dimensional Gaussian Hilbert space generated by a standard Gaussian random variable ξ with mean 0 and variance 1. In this case

$$\mathcal{P}_n(\mathfrak{X}) = \mathrm{span}_{\mathbb{R}} \left\{ H_k(\xi);\ \ k \leq n \right\}.$$

Since $\mathbb{E}[H_j(\xi) H_k(\xi)] = 0$ for $j \neq k$, we deduce that

$$\mathfrak{X}^{:n:} = \mathrm{span} \left\{ H_n(\xi) \right\}.$$

Moreover, (3.1.10) implies that, $\forall n \geq 0$ we have

$$\xi^n = \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}\big[\, \xi^{n-k} \,\big] H_k(\xi) = \sum_{j=0}^{\lfloor n/2 \rfloor} \binom{n}{2j} \mathbb{E}\big[\, \xi^{2j} \,\big] H_{n-2j}(\xi) \tag{3.1.23}$$

$$\stackrel{(1.1.9)}{=} \sum_{j=0}^{\lfloor n/2 \rfloor} (2j-1)!! \binom{n}{2j} H_{n-2j}(\xi).$$

In particular,

$$\mathrm{Proj}_n(\xi^n) = H_n(\xi). \tag{3.1.24}$$

If $\mathfrak{X}$ is a separable Gaussian Hilbert space and $\underline{X} = (X_n)_{n \geq 1}$ is a complete orthonormal basis of $\mathfrak{X}$, then the computations at the end of Subsection 3.1.2 show that the collection $H_\alpha(X_1, \dots, X_m)$, $m \in \mathbb{N}$, $\alpha \in \mathbb{N}_0^m$, is an orthogonal basis of $\mathcal{F}(\mathfrak{X})$ □

3.1.5. Wick products and the diagram formula. Fix a probability space $(\Omega, \mathcal{S}, \mathbb{P})$ and a separable real Gaussian Hilbert space $\mathfrak{X} \subset L^2(\Omega, \mathcal{S}, \mathbb{P})$. Denote by $\mathcal{F}(\mathfrak{X})$ the Wiener chaos of $\mathfrak{X}$, and consider the Wiener chaos decomposition

$$\mathcal{F}(\mathfrak{X}) = \widehat{\bigoplus_{n\geq 0}} \mathfrak{X}^{:n:}.$$

We have bilinear maps

$$\mathfrak{X}^{:m:} \times \mathfrak{X}^{:n:} \to \mathfrak{X}^{:(m+n):}, \quad \mathfrak{X}^{:m:} \times \mathfrak{X}^{:n:} \ni (\xi, \eta) \mapsto \xi \bullet \eta := \operatorname{Proj}_{m+n}(\xi\eta).$$

Remark 3.1.17. If $\underline{X} = (X_k)_{k\geq 1}$ is a complete orthonormal basis of $\mathfrak{X}$, and $\alpha, \beta \in \mathbb{N}_0$ are such that $|\alpha| = m$, $|\beta| = n$, then

$$H_\alpha(\underline{X}) \bullet H_\beta(\underline{X}) = H_{\alpha+\beta}(\underline{X}). \tag{3.1.25}$$

Indeed,

$$H_\alpha(\underline{X}) \bullet H_\beta(\underline{X}) = \sum_{|\gamma|=m} c_\gamma H_\gamma(\underline{X}).$$

Now observe that for any multi-index γ such that $|\gamma| = m + n$, and $\gamma \neq \alpha + \beta$ the coefficient of $\underline{X}^{\alpha+\beta}$ in $H_\gamma(\underline{X})$ is 0, while the coefficient of $\underline{X}^{\alpha+\beta}$ in $H_{\alpha+\beta}(\underline{X})$ is 1. □

Definition 3.1.18. Fix a polynomial $P \in \mathbb{R}[x_1, \dots, x_n]$ of degree m. For any $\xi_1, \dots, \xi_n \in \mathfrak{X}$, the random variable $\operatorname{Proj}_m P(\xi_1 \cdots \xi_n) \in \mathfrak{X}^{:m:}$ is called the *Wick polynomial* associated to $P(\xi_1, \dots, \xi_n)$ and it is denoted by $: P(\xi_1 \cdots \xi_n) :$. □

Theorem 3.1.19. *Let $\mathfrak{X}$ be a separable Gaussian Hilber space and $\xi_1, \dots, \xi_n \in \mathfrak{X}$. Then*

$$: \xi_1 \cdots \xi_n : \; = \sum_\Gamma (-1)^{r(\Gamma)} w(\Gamma), \tag{3.1.26}$$

where the summation is over all the Feynman diagrams with vertices labelled by $\xi_1, \dots, \xi_n$.

Proof. Denote by $L(\xi_1, \dots, \xi_n)$ the left-hand-side of (3.1.26) and by $R(\xi_1, \dots, \xi_n)$ its right-hand side. Observe that both L and R are symmetric, multi-linear forms in the variables $\xi_1, \dots, \xi_m$ and thus

$$L(\xi_1, \dots, \xi_n) = R(\xi_1, \dots, \xi_n), \quad \forall \xi_1, \dots, \xi_n \Longleftrightarrow L(\underbrace{\xi, \dots, \xi}_{n}) = R(\underbrace{\xi, \dots, \xi}_{n}), \;\; \forall \xi \in \mathfrak{X}, \; \operatorname{Var}\big[\,\xi\,\big] = 1.$$

Let $\xi \in \mathfrak{X}$ such that $\operatorname{Var}\big[\,\xi\,\big] = 1$. Then

$$L(\underbrace{\xi, \dots, \xi}_{n}) =: \xi^n : \overset{(3.1.24)}{=} H_n(\xi).$$

Then

$$R(\underbrace{\xi, \dots, \xi}_{n}) = \sum_{r\geq 0} (-1)^r \sum_{\gamma \in \operatorname{Feyn}^r(n)} \xi^{m-2r} \overset{(3.1.19)}{=} H_n(\xi).$$

Hence

$$R(\underbrace{\xi, \dots, \xi}_{n}) = L(\underbrace{\xi, \dots, \xi}_{n}).$$

□

Corollary 3.1.20. *Suppose that* $\xi_1, \dots, \xi_n \in \mathfrak{X}$ *is an orthonormal system, i.e.,*

$$\mathbb{E}\big[\,\xi_i\xi_j\,\big] = \delta_{ij}, \ \ \forall i, j.$$

Then for any $\alpha \in \mathbb{N}_0^n$ *we have*

$$:\xi^\alpha : = :\xi_1^{\alpha_1}\cdots\xi_n^{\alpha_n} := H_\alpha(\xi_1, \dots, \xi_n).$$

Proof. Set $m := |\alpha|$. Since the random variables $\xi_1, \dots, \xi_n$ are independent, we deduce that

$$\mathbb{E}[\,H_\beta(\xi_1, \dots, \xi_n)H_\gamma(\xi_1, \dots, \xi_n)\,] = \prod_{j=1}^n \mathbb{E}[\,H_{\beta_j}(\xi_j)H_{\gamma_j}(\xi_j)\,], \ \ \forall \beta, \gamma \in \mathbb{Z}^n_{\geq 0}.$$

We deduce from the orthogonality of the Hermite polynomials that the collection

$$\big(\,H_\beta(\xi_1, \dots, \xi_n)\,\big)_{|\beta|\leq m}$$

is an orthogonal basis of $\mathcal{P}_m(\xi_1, \dots, \xi_n)$. In particular, we have a unique linear decomposition

$$\xi^\alpha = \sum_{|\beta|\leq m} c_\beta H_\beta(\xi), \tag{3.1.27a}$$

$$:\xi^\alpha : = \sum_{|\beta|=m} c_\beta H_\beta(\xi). \tag{3.1.27b}$$

For any multi-index β such that $|\beta| = m$, the coefficient of ξ^β in the right-hand-side of (3.1.27a) is c_β. We deduce that $c_\beta = 0$ for all β such that $|\beta| = m$ and $\beta \neq \alpha$. The conclusion of Corollary 3.1.20 is now obvious. □

Corollary 3.1.21. *The space*

$$\operatorname{span}\big\{\,:\xi_1\xi_2\cdots\xi_n:\,;\ \ \xi_1, \dots, \xi_n \in \mathfrak{X}\,\big\}$$

is dense in $\mathfrak{X}^{:n:}$.

Proof. Follows from Example 3.1.16 and Corollary 3.1.20. □

Corollary 3.1.22. *Suppose that* $\underline{X} = (X_k)_{k\geq 1}$ *is a complete orthonormal basis of* $\mathfrak{X}$. *Then the collection*

$$H_\alpha\big(\,\underline{X}\,\big), \ \ \alpha \in \mathbb{N}_0^{\mathbb{N}}, \ \ |\alpha| = n,$$

is an complete orthogonal basis of $\mathfrak{X}^{:n:}$.

Proof. Let $\xi_1, \dots, \xi_n \in \mathfrak{X}$. Then

$$\xi_1\cdots\xi_n = \sum_{\substack{\alpha\in\mathbb{N}_0^n\\ |\alpha|=n}} c_\alpha X^\alpha, \ \ X^\alpha = \prod_{n\mathbb{N}} X_n^{\alpha_n},$$

where the above series converges in L^2. We deduce

$$:\xi_1\cdots\xi_n := \sum_{\substack{\alpha\in\mathbb{N}_0^n\\ |\alpha|=n}} c_\alpha H_\alpha\big(\,\underline{X}\,\big).$$

□

Theorem 3.1.23 (Diagram Formula)**.** *Consider an array of random variables*

$$\mathcal{A} = \big\{ \xi_{ij} \in \mathfrak{X};\ \ 1 \le i \le k,\ \ 1 \le j \le \ell_i \big\}.$$

Denote by $\mathrm{Feyn}\big[\mathcal{A}\big] \subset \mathrm{Feyn}(\mathcal{A})$ *the collection of Feynman diagrams with vertices in* $\mathcal{A}$ *and* compatible with the array structure *of* $\mathcal{A}$*. This means that no edge connects vertices situated on the same row of the array* $\mathcal{A}$*. Denote by* $\mathrm{Feyn}^*\big[\mathcal{A}\big]$ *the subcollection of* $\mathrm{Feyn}\big[\mathcal{A}\big]$ *consisting of complete diagrams. For* $i = 1, \dots, k$ *we set*

$$Y_i := \mathrm{Proj}_{\ell_i}\left(\prod_{j=1}^{\ell_i} \xi_{ij} \right),\ \ \ell = \ell_1 + \cdots + \ell_k.$$

In other words, Y_i *is the Wick product of the variables situated on the* i*-th row. Then*

$$\mathbb{E}\big[\, Y_1 \cdots Y_k \,\big] = \sum_{\Gamma \in \mathrm{Feyn}^*\left[\mathcal{A}\right]} w(\Gamma), \tag{3.1.28a}$$

$$Y_1 \cdots Y_k = \sum_{\Gamma \in \mathrm{Feyn}\left[\mathcal{A}\right]} : w(\Gamma) := \sum_{\Gamma \in \mathrm{Feyn}\left[\mathcal{A}\right]} \mathrm{Proj}_{\ell - 2r(\Gamma)}\, w(\Gamma). \tag{3.1.28b}$$

Proof. I follow the approach in [**76**, Thm. 3.12, 3.15]. Denote by $\mathcal{A}^i$ the i-th row of the array $\mathcal{A}$ and by $\mathrm{Feyn}^*(\mathcal{A})$ the collection of all complete Feynman diagrams with vertices in the array, *not necessarily compatible with the array.* I want to emphasize that $\mathrm{Feyn}^*\big[\mathcal{A}\big] \subset \mathrm{Feyn}^*(\mathcal{A})$.

Theorem 3.1.19 implies that

$$Y_i = \sum_{\Gamma_i \in \mathrm{Feyn}(\mathcal{A}^i)} (-1)^{r(\Gamma_i)} w(\Gamma_i),$$

$$Y_1 \cdots Y_k = \prod_{i=1}^{k} \left(\sum_{\Gamma_i \in \mathrm{Feyn}(\mathcal{A}^i)} (-1)^{r(\Gamma_i)} w(\Gamma_i) \right)$$

$$= \sum_{(\Gamma_1, \dots, \Gamma_k) \in \mathrm{Feyn}(\mathcal{A}^1) \times \cdots \times \mathrm{Feyn}(\mathcal{A}^k)} (-1)^{\sum_{i=1}^k r(\Gamma_i)} \prod_{i=1}^{k} w(\Gamma_i),$$

so that

$$\mathbb{E}\big[\, Y_1 \cdots Y_k \,\big] = \sum_{(\Gamma_1, \dots, \Gamma_k) \in \mathrm{Feyn}(\mathcal{A}^1) \times \cdots \times \mathrm{Feyn}(\mathcal{A}^k)} (-1)^{\sum_{i=1}^k r(\Gamma_i)} \mathbb{E}\left[\prod_{i=1}^{k} w(\Gamma_i) \right].$$

Given $(\Gamma_1, \dots, \Gamma_k) \in \mathrm{Feyn}(\mathcal{A}^1) \times \cdots \times \mathrm{Feyn}(\mathcal{A}^k)$ we denote by $\mathrm{Feyn}^*(\Gamma_1, \dots, \Gamma_k)$ the subcollection of $\mathrm{Feyn}^*\big(\mathcal{A}\big)$ consisting of diagrams that contain $\Gamma_1 \cup \cdots \cup \Gamma_k$ as a subdiagram.

We deduce from Wick's formula (3.1.20) that

$$\mathbb{E}\left[\prod_{i=1}^{k} w(\Gamma_i) \right] = \sum_{\Gamma' \in \mathrm{Feyn}^*(\Gamma_1, \dots, \Gamma_k)} w(\Gamma').$$

Hence

$$\mathbb{E}\big[\, Y_1 \cdots Y_k \,\big] = \sum_{\Gamma' \in \mathrm{Feyn}^*(\Gamma_1,\dots,\Gamma_k)} w(\Gamma') \underbrace{\left(\sum_{\substack{(\Gamma_1,\dots,\Gamma_k)\in \mathrm{Feyn}(\mathcal{A}^1)\times\cdots\times\mathrm{Feyn}(\mathcal{A}^k)\\ \Gamma_1\cup\cdots\cup\Gamma_k\subset\Gamma'}} (-1)^{\sum_{i=1}^k r(\Gamma_i)} \right)}_{=:S(\Gamma')}.$$

For any $\Gamma' \in \mathrm{Feyn}^*(\Gamma_1,\dots,\Gamma_k)$ we have

$$S(\Gamma') = \prod_{i=1}^k \underbrace{\left(\sum_{\substack{\Gamma_i\in\mathrm{Feyn}(\mathcal{A}^i)\\ \Gamma_i\subset\Gamma'}} (-1)^{r(\Gamma_i)} \right)}_{=:S_i(\Gamma')}.$$

Given $\Gamma' \in \mathrm{Feyn}^*(\mathcal{A})$ we denote by $\Gamma' \cap \mathcal{A}^i$ the subdiagram of Γ' consisting only of edges connecting vertices on the i-th row of $\mathcal{A}$. We have

$$S_i(\Gamma') = \sum_{\Gamma_i \subset \Gamma' \cap \mathcal{A}^i} (-1)^{r(\Gamma_i)}.$$

Now observe that $S_i(\Gamma') = 0$ if $\Gamma' \cap \mathcal{A}^i \neq \emptyset$ and it is $= 1$ otherwise. Indeed, if $r = r(\Gamma' \cap \mathcal{A}^i) > 0$ then

$$\sum_{\Gamma_i\subset\Gamma'\cap\mathcal{A}^i} (-1)^{r(\Gamma_i)} = \sum_{S\subset\mathbb{I}_r} (-1)^{|S|} = \sum_{j=0}^r (-1)^j \binom{r}{j} = 0.$$

Thus

$$S(\Gamma') = \begin{cases} 1, & \Gamma' \in \mathrm{Feyn}^*\big[\,\mathcal{A}\,\big],\\ 0, & \Gamma' \in \mathrm{Feyn}^*(\mathcal{A}) \setminus \mathrm{Feyn}^*\big[\,\mathcal{A}\,\big]. \end{cases}$$

This proves (3.1.28a).

Denote by L, respectively R the left-hand-side respectively the right-hand-side of the equality (3.1.28b). For any random variables

$$\eta_1,\dots,\eta_m \in \mathrm{span}\big\{\xi_{ij} \in \mathfrak{X},;\ \ 1 \le i \le k,\ \ 1 \le j \le \ell_i\,\big\}$$

we denote by $\mathcal{A}_\eta$ the array obtained from $\mathcal{A}$ by adding an extra row consisting of the variables $\eta_1,\dots,\eta_m$. Set $Z := (:\eta_1\cdots\eta_m:)$. Then (3.1.28a) applied to $\mathcal{A}_\eta$ implies that we have

$$\mathbb{E}[LZ] = \mathbb{E}[RZ] \Longleftrightarrow \mathbb{E}\big[\,(L-R)Z\,\big] = 0.$$

The equality (3.1.28b) now follows from Corollary 3.1.21. □

Example 3.1.24. Let us apply the diagram formula in the special case when the array $\mathcal{A} = (\xi_{ij})$ consists of two rows of lengths $\ell_1 \le \ell_2$ and and $\xi_{ij} = \xi$, $\forall i = 1,2, j = 1,\dots,\ell_i$, $\mathbb{E}\big[\,\xi^2\,\big] = 1$. Then

$$Y_i = H_{\ell_i}(\xi)$$

and we deduce

$$H_{\ell_1}(\xi)H_{\ell_2}(\xi) = \sum_{\Gamma\in\mathrm{Feyn}\left[\,\mathcal{A}\,\right]} H_{\ell_1+\ell_2-2r(\Gamma)}(\xi) = \sum_{r=0}^{\ell_2} r!\binom{\ell_1}{r}\binom{\ell_2}{r} H_{\ell_1+\ell_2-2r}(\xi).$$

More generally, assume the array has two rows, but the variables on the first row are equal to ξ_1, while the variables on the second row are equal to ξ_2, $\mathbb{E}\big[\,\xi_1^2\,\big]=\mathbb{E}\big[\,\xi_2^2\,\big]=1$. If $c:=\mathbb{E}[\xi_1\xi_2]$, then

$$H_{\ell_1}(\xi_1)H_{\ell_2}(\xi_2)=\sum_{r=0}^{\ell} r!\binom{\ell_1}{r}\binom{\ell_2}{r}c^r\,\mathrm{Proj}_{\ell_1+\ell_2-2r}(\xi_1^{\ell_1-r}\xi_2^{\ell_2-r}). \tag{3.1.29}$$

If $\ell_1=\ell_2=\ell$, then (3.1.28a) implies that

$$\mathbb{E}\big[\,H_{\ell_1}(\xi_1)H_{\ell_2}(\xi_2)\,\big]=\ell!\binom{2\ell}{\ell}c^\ell. \tag{3.1.30}$$

□

The equality (3.1.28b) implies[2] that for any positive integer n there exists a constant $C(n)>0$ such that for any $X\in\mathcal{P}_n(\mathfrak{X})$ we have

$$\|X\|_{L^4}\leq C(n)\|X\|_{L^2}.$$

In particular, this shows that the bilinear map

$$\mathfrak{X}^{:m:}\times\mathfrak{X}^{:n:}\ni(X,Y)\mapsto X\bullet Y:=\mathrm{Proj}_{m+n}(XY)\in\mathfrak{X}^{:m+n:}$$

is continuous. Corollary 3.1.22 now implies that the multiplication $\bullet$ satisfies the associativity property

$$(\xi\bullet\eta)\bullet\zeta=\xi\bullet(\eta\bullet\zeta),\;\;\forall\xi\in\mathfrak{X}^{:\ell:},\;\eta\in\mathfrak{X}^{:m:},\;\zeta\in\mathfrak{X}^{:n:},\;\;\forall\ell,m,n\in\mathbb{N}_0. \tag{3.1.31}$$

Indeed, (3.1.25) shows that the above equality is true for

$$\xi,\eta,\zeta\in\big\{\,H_\alpha(\underline{X});\;\;\alpha\in\mathbb{Z}_{\geq 0}^{\mathbb{N}},\;\;|\alpha|<\infty\,\big\}.$$

The general case follows from the multi-linearity and continuity of (3.1.31) in ξ,η,ζ. A similar argument shows that

$$\xi\bullet\eta=\eta\bullet\xi,\;\;\forall\xi\in\mathfrak{X}^{:m:},\;\;\eta\in\mathfrak{X}^{:n:}. \tag{3.1.32}$$

We thus obtain a structure of commutative and associative $\mathbb{R}$-algebra on $\mathfrak{X}$ called the *Wick algebra* of $\mathfrak{X}$. The product $\bullet$ is called the *Wick product*. Note that for $\xi_1,\dots,\xi_m\in\mathfrak{X}$ we have

$$\xi_1\bullet\cdots\bullet\xi_m=:\xi_1\cdots\xi_m:$$

In general, if

$$\xi=\sum_{n\geq 0}\xi_n,\;\;\eta=\sum_{n\geq 0}\eta_n,\;\;\xi_n,\eta_n\in\mathfrak{X}^{:n:},$$

then

$$\xi\bullet\eta:=\sum_{n\geq 0}\left(\sum_{j+k=n}\xi_j\bullet\eta_k\right).$$

Example 3.1.25 (The Wick exponential)**.** Suppose that $X\in\mathfrak{X}$ and $v=\mathrm{Var}\big[\,X\,\big]$. Define the *Wick exponential*

$$:e^X:=\sum_{n\geq 0}\frac{1}{n!}:X^n:.$$

[2]See [**76**, Lemma 3.44] for details.

Using (3.1.24) we deduce

$$:X^n := v^{n/2} H_n(v^{-1/2}X) \stackrel{(3.1.13)}{=} H_n(x|v).$$

The equality (3.1.12) implies

$$:e^X := e^{X-v/2} = e^{X-\frac{1}{2}\operatorname{Var}[X]}. \tag{3.1.33}$$

If $X, Y \in \mathfrak{X}$, then

$$\big(\,:e^X:\,\big)\big(\,:e^Y:\,\big) = e^{X+Y-\frac{1}{2}(\operatorname{Var}[X]+\operatorname{Var}[Y])}$$

Then

$$\mathbb{E}\big[\,\big(\,:e^X:\,\big)\big(\,:e^Y:\,\big)\,\big] = \mathbb{E}\big[\,e^{X+Y}\,\big]e^{-\frac{1}{2}(\operatorname{Var}[X]+\operatorname{Var}[Y])}$$

($X+Y$ is centered Gaussian)

$$\stackrel{(1.1.7)}{=} e^{\frac{1}{2}\big(\operatorname{Var}[X+Y]-\operatorname{Var}[X]-\operatorname{Var}[Y]\big)} = e^{\mathbb{E}[XY]}.$$

□

3.1.6. Fock spaces. We have shown that the Wiener chaos is a Hilbert space equipped with a structure commutative and associative algebra. In this subsection we will describe a general procedure that associates to an abstract separable Hilbert space H, a bigger space equipped with a structure of commutative and associative algebra. This bigger space is called the Fock space of H and plays an important role in quantum field theory, [**151**, Chap.3]. We will show that the Fock space of a Gaussian Hilbert space is naturally isomorphic as a Hilbert space and as algebra with the Wiener chaos.

The construction of the Fock spaces is based on the tensor product of two separable Hilbert spaces H_1, H_2 defined as follows.

Construct first the *algebraic* tensor product $H_1 \otimes H_2$. The universality property of the tensor product implies that there exists a unique inner product $(-,-)_{H_1\otimes H_2}$ of $H_1 \otimes H_2$ such that, for any $x_i, y_i \in H_i$, $i = 1, 2$, we have

$$\big(\,x_1 \otimes x_2\,,\,y_1 \otimes y_2\,\big)_{H_1\otimes H_2} = (x_1, y_1)_{H_1} \cdot (x_2, y_2)_{H_2}.$$

We denote by $H_1 \hat{\otimes} H_2$ the completion of $H_1 \otimes H_2$ with respect to the norm defined by the above inner product. The Hilbert space $H_1 \hat{\otimes} H_2$ is called the (analytic) *tensor product* of the Hilbert spaces H_1, H_2.

Theorem 3.1.26. *Suppose that $(M_j, \mathcal{M}_j, \mu_j)$, $i = 1, 2$, are two measured spaces such that the Hilbert spaces $H_j = L^2(M_j, \mathcal{M}_j \mu_j)$ are separable.*[3] *Then there exists a unique isomorphism of Hilbert space*

$$H_1 \hat{\otimes} H_2 \to L^2\big(\,M_1 \times M_2, \mu_1 \otimes \mu_2\,\big),$$

such that

$$f_1 \hat{\otimes} f_2 \mapsto \big(\,f_1 \boxtimes f_2 : M_1 \times M_2 \to \mathbb{R}\,\big),\ \ f_1 \boxtimes f_2(x_1, x_2) = f_1(x_1) f_2(x_2).$$

□

[3] E.g., this happens when the sigma-algebras $\mathcal{M}_j$ are countably generated and the measures μ_i are sigma-finite.

For a proof and more details we refer to [**134**, Thm. II.10], or the original source, [**106**].

The tensor product of two separable Hilbert spaces H_1, H_2 can also be realized as the space of Hilbert-Schmidt bilinear functionals $u : H_1 \times H_2 \to \mathbb{R}$. This means that for any complete orthonormal bases $(e_m)_{m\geq 1}$ of H_1 and $(f_n)_{n\geq 1}$ of H_2 we have

$$\sum_{m,n\geq 1} |u(e_m, f_n)|^2 < \infty.$$

The tensor product of Hilbert spaces enjoys the usual commutativity and associativity properties

$$H_1 \hat{\otimes} H_2 \cong H_2 \hat{\otimes} H_1, \;\; (H_1 \hat{\otimes} H_2)\hat{\otimes} H_3 \cong H_1 \hat{\otimes}(H_2 \hat{\otimes} H_3).$$

Given a separable Hilbert space H we denote by $H^{\odot n}$ its *algebraic* n-th symmetric product, i.e., the subspace of $H^{\otimes n}$ consisting of elements fixed by the obvious action of the symmetric group $\mathfrak{S}_n$. The closure of $H^{\odot n}$ in $H^{\hat{\otimes} n}$ is denoted by $H^{\hat{\odot} n}$ and it is called the analytic *n-th symmetric power* of H.

Note that we have a natural projector $\mathbf{Sym} : H^{\otimes n} \to H^{\odot n}$ defined by

$$\mathbf{Sym}[x_1 \otimes \cdots \otimes x_n] := \frac{1}{n!} \sum_{\sigma\in\mathfrak{S}_n} x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}, \;\; \forall x_1, \dots, x_n \in H.$$

For $x_1, \dots, x_n \in H$ we set

$$x_1 \odot \cdots \odot x_n := \sqrt{n!}\,\mathbf{Sym}[x_1 \otimes \cdots \otimes x_n] = \frac{1}{\sqrt{n!}} \sum_{\sigma\in\mathfrak{S}_n} x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}. \tag{3.1.34}$$

Note that

$$x^{\odot n} = \underbrace{x \odot \cdots \odot x} = \sqrt{n!} x^{\otimes n}$$

and

$$\|x^{\odot n}\|^2 = n!\|x\|^{2n}. \tag{3.1.35}$$

This is a manifestation of a more general phenomenon.

Lemma 3.1.27. *If $e_1, \dots, e_n$ is an orthonormal system in H and $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$. We set*

$$e^{\odot\alpha} := \underbrace{e_1 \odot \cdots \odot e_1}_{\alpha_1} \odot \cdots \odot \underbrace{e_n \odot \cdots \odot e_n}_{\alpha_n} = e_1^{\odot\alpha_1} \odot \cdots \odot e_n^{\odot\alpha_n}.$$

Then

$$\|e^{\odot\alpha}\|^2 := \|e_1^{\odot\alpha_1} \odot \cdots \odot e_n^{\odot\alpha_n}\|^2 = \alpha! = (\alpha_1!) \cdots (\alpha_n!) = \|H_\alpha\|^2. \tag{3.1.36}$$

Proof. For any multi-index $\alpha \in \mathbb{N}^n$ we define an α-coloring, or a coloring of type α, to be a map $\pi : \mathbb{I}_{|\alpha|} \to \mathbb{I}_n$ such that

$$\#\pi^{-1}(k) = \alpha_k, \;\; \forall k = 1, \dots, n.$$

Think of π as defining a coloring of $|\alpha|$ objects with n colors $e_1, \dots, e_n$ so that exactly α_k objects have color e_k. We denote by $\mathcal{P}_\alpha$ the set α-colorings. Hence

$$\#\mathcal{P}_\alpha = \binom{|\alpha|}{\alpha_1, \dots, \alpha_n} := \frac{|\alpha|!}{\alpha_1! \cdots \alpha_n!}.$$

The symmetric group $\mathfrak{S}_{|\alpha|}$ acts transitively on $\mathcal{P}_\alpha$

$$\mathcal{P}_\alpha \ni \times \mathfrak{S}_{|\alpha|} \ni (\pi, \sigma) \mapsto \pi \circ \sigma \in \mathcal{P}_\alpha$$

and the stabilizer of an α-covering π is isomorphic to

$$\mathfrak{S}_\alpha := \mathfrak{S}_{\alpha_1} \times \cdots \times \mathfrak{S}_{\alpha_n}.$$

To each $\pi \in \mathcal{P}_\alpha(E)$ we associate the element

$$e_\pi := e_{\pi(1)} \otimes \cdots \otimes e_{\pi(|\alpha|)} \in H^{\otimes |\alpha|}.$$

Let us observe that

$$(e_\pi, e_{\pi'}) = \prod_{k=1}^{|\alpha|} \big(e_{\pi(k)}, e_{\pi'(k)} \big)_H = \delta_{\pi\pi'}, \;\; \forall \pi, \pi' \in \mathcal{P}_\alpha.$$

Then

$$e^{\odot \alpha} = \frac{\prod_{k=1}^n \alpha_k!}{\sqrt{|\alpha|!}} \sum_{\pi \in \mathcal{P}_\alpha} e_\pi,$$

so that

$$\|e^{\odot\alpha}\|^2 = \left(\frac{\prod_{k=1}^n \alpha_k!}{\sqrt{|\alpha|!}} \right)^2 \binom{|\alpha|}{\alpha_1, \dots, \alpha_n} = (\alpha_1!) \cdots (\alpha_n!).$$

□

The above lemma implies[4] that we have continuous bilinear map

$$\odot : H^{\odot m} \times H^{\odot n} \to H^{\odot(m+n)}$$

defined by

$$X \odot Y := \sqrt{\binom{n+m}{m}} \, \mathbf{Sym}\big(X \otimes Y \big)$$
$$= \sqrt{\binom{n+m}{m}} \, \mathbf{Sym}\big(\mathbf{Sym}(X) \otimes \mathbf{Sym}(Y) \big),$$

$\forall X \in H^{\odot m}, \;\; Y \in H^{\odot n}$.

Proposition 3.1.28. *Let $n_1, n_2, n_3 \in \mathbb{N}$ and $X_i \in H^{\odot n_i}$, $i = 1, 2, 3$. Then*

$$\big(X_1 \odot X_2 \big) \odot X_3 = X_1 \odot \big(X_2 \odot X_3 \big).$$

Proof. Set $\alpha = (n_1, n_2, n_3)$, $n = n_1 + n_2 + n_3$. We can assume that

$$H = L^2(I) = L^2(I, \boldsymbol{\lambda}), \;\; I = [0, 1].$$

Then

$$X_i \in L^2(I^{n_i}) = L^2\big(I^{n_i}, \boldsymbol{\lambda}^{\otimes n_i} \big) \cong H^{\otimes n_i}.$$

The symmetric tensor product $H^{\odot n_i}$ is isomorphic as a vector space with the subspace of symmetric functions in $L^2\big(I^{n_i} \big)$.

The symmetric group $\mathfrak{S}_n$ acts on I^n by permuting the coordinates. Thus, for

$$\boldsymbol{t} := (t_1, \dots, t_n), \;\; \sigma \in \mathfrak{S}_n$$

[4]The details are straightforward and not particularly illuminating.

we set

$$\sigma \cdot \boldsymbol{t} := \big(t_{\sigma(1)}, \dots, t_{\sigma(n)} \big)$$

For $F \in L^2(I^m)$ we have

$$\mathbf{Sym}(F)(\boldsymbol{t}) = \frac{1}{m!} \sum_{\sigma \in \mathfrak{S}_m} F(\sigma \cdot \boldsymbol{t}).$$

We set $G := \mathfrak{S}_n$. We introduce the following notation.

- G_1 is the subgroup of G that permutes only the first n_1 variables in $(t_1, \dots, t_n)$.
- G_3 is the subgroup of G that permutes only the last n_3 variables in $(t_1, \dots, t_n)$.
- G_2 is the subgroup of G that permutes only the middle n_2 variables in $(t_1, \dots, t_n)$.
- $G_{1,2}$ is the subgroup of G that permutes only the first n_1+n_2 variables in $(t_1, \dots, t_n)$
- $G_{2,3}$ is the subgroup of G that permutes only the last n_2+n_3 variables in $(t_1, \dots, t_n)$.

We have inclusions

$$H_L := G_{1,2} \times G_3 \subset G, \;\; H_R := G_1 \times G_{2,3} \subset G.$$

For any subgroup $H \subset G$ we denote by $H \backslash G$ the set or orbits of the left-hand action of H on G. We have the following commutative diagram of surjective maps.

$$\begin{array}{ccccc} & \overset{L_1}{\swarrow} & G & \overset{R_1}{\searrow} & \\ H_L \backslash G & & \big\downarrow p & & H_R \backslash G \\ & \underset{L_2}{\searrow} & * & \underset{R_2}{\swarrow} & \end{array} \tag{3.1.37}$$

Above $\{*\}$ denotes a generic singleton, and L_1, R_1 are the canonical projections onto the corresponding spaces of orbits.

For any surjective map $\Phi : A \to B$, A, B finite sets, and any $F : A \to \mathbb{R}$ we denote by $\Phi_*(F)$ its summation along the fibers of Φ. More explicitly, $\Phi_*(F)$ is the function $B \to \mathbb{R}$ defined by

$$\Phi_*(F)(b) = \sum_{a \in \Phi^{-1}(b)} F(a).$$

For each $\boldsymbol{t} \in I^n$ we have a function

$$F_{\boldsymbol{t}} : G \to \mathbb{R}, \;\; F_{\boldsymbol{t}}(g) = \frac{1}{n!} X_1 \boxtimes X_2 \boxtimes X_3 \big(g \cdot \boldsymbol{t} \big).$$

Note that

$$\mathbf{Sym} \big(X_1 \boxtimes X_2 \boxtimes X_3 \big)(\boldsymbol{t}) = \frac{1}{n!} \sum_{g \in G} F_{\boldsymbol{t}}(g) = p_*(F_{\boldsymbol{t}})$$

Let $H_L \cdot g \in H_L \backslash G$. Then

$$(L_1)_* F_{\boldsymbol{t}}(H_L \cdot g) = \frac{1}{n!} \sum_{h \in H_L} X_1 \boxtimes X_2 \boxtimes X_3 \big(hg \cdot \boldsymbol{t} \big)$$

$$= \frac{(n_1+n_2)! n_3!}{n!} \mathbf{Sym}(X_1 \boxtimes X_2) \boxtimes X_3(g\boldsymbol{t})$$

The colorings of type $\alpha = (n_1 + n_2, n_3)$ are maps $\pi : \mathbb{I}_n \to \{1, 2\}$ such that

$$\#\pi^{-1}(1) = n_1 + n_2 \text{ and } \#\pi^{-1}(2) = n_3.$$

The group G acts on the space $\mathcal{P}_\alpha$

$$\mathcal{P}_\alpha \times G \ni (\pi, g) \mapsto \pi \circ g, \quad \forall \pi \in \mathcal{P}_\alpha, \quad g \in G = \mathfrak{S}_n.$$

We obtain a bijection $H_L \backslash H \to \mathcal{P}_\alpha$

$$H_L \cdot g \mapsto \pi_0 \circ g$$

where $\pi_0 \in \mathcal{P}_\alpha$ is defined by

$$\pi_0(i) = \begin{cases} 1, & i \leq n_1 + n_2, \\ 2, & n_1 + n_2 < i \leq n. \end{cases}$$

For $\pi \in \mathcal{P}_\alpha$ we set

$$\pi \cdot \boldsymbol{t} = (t_{\pi^{-1}(1)}, t_{\pi^{-1}(2)})$$

where for any subset $J = \{j_1, \dots, j_k\} \subset \mathbb{I}_n$, $j_1 < \cdots < j_k$, we set

$$t_J = (t_{j_1}, \dots, t_{j_k}).$$

We deduce that

$$p_*(F_{\boldsymbol{t}}) = (L_2)_*(L_1)_* F_{\boldsymbol{t}} = \frac{(n_1 + n_2)! n!}{n!} \sum_{\pi \in \mathcal{P}_\alpha} \mathbf{Sym}(X_1 \boxtimes X_2) \boxtimes X_3\big(\pi \cdot \boldsymbol{t}\big)$$

$$= \frac{1}{\#\mathcal{P}_\alpha} \sum_{\pi \in \mathcal{P}_\alpha} \mathbf{Sym}(X_1 \boxtimes X_2) \boxtimes X_3\big(\pi \cdot \boldsymbol{t}\big) = \mathbf{Sym}\big(\mathbf{Sym}(X_1 \boxtimes X_2) \boxtimes X_3\big)(\boldsymbol{t}).$$

Since $p_* = (L_2 \circ L_1)* = (L_2)_* \circ (L_1)_*$, we deduce

$$\mathbf{Sym}\big(X_1 \boxtimes X_2 \boxtimes X_3\big)(\boldsymbol{t}) = \mathbf{Sym}\big(\mathbf{Sym}(X_1 \boxtimes X_2) \boxtimes X_3\big)(\boldsymbol{t})$$

Using the right-hand-side of the diagram (3.1.37) we deduce in a similar fashion that

$$\mathbf{Sym}\big(X_1 \boxtimes X_2 \boxtimes X_3\big)(\boldsymbol{t}) = \mathbf{Sym}\big(X_1 \boxtimes \mathbf{Sym}(X_2 \boxtimes X_3)\big)(\boldsymbol{t}).$$

Hence

$$\mathbf{Sym}\big(\mathbf{Sym}(X_1 \boxtimes X_2) \boxtimes X_3\big) = \mathbf{Sym}\big(X_1 \boxtimes \mathbf{Sym}(X_2 \boxtimes X_3)\big)$$

On the other hand,

$$\mathbf{Sym}(X_1 \boxtimes X_2) = \binom{n_1 + n_2}{n_1}^{-1/2} X_1 \odot X_2$$

$$\mathbf{Sym}\big(\mathbf{Sym}(X_1 \boxtimes X_2) \boxtimes X_3\big) = \binom{n}{n_1 + n_2}^{-1/2} \binom{n_1 + n_2}{n_1}^{-1/2} (X_1 \odot X_2) \odot X_3$$

$$= \binom{n}{n_1, n_2, n_3}^{-1/2} (X_1 \odot X_2) \odot X_3.$$

Similarly

$$\mathbf{Sym}\big(X_1 \boxtimes \mathbf{Sym}(X_2 \boxtimes X_3)\big) = \binom{n}{n_1, n_2, n_3}^{-1/2} X_1 \odot (X_2 \odot X_3).$$

□

We obtain in this fashion a graded associative and commutative algebra

$$\bigoplus_{n\geq 0} H^{\hat{\odot} n}.$$

Its completion

$$\widehat{\bigoplus_{n\geq 0}} H^{\hat{\odot} n}$$

is called the *Fock space* of H and it is denoted by $\boldsymbol{F}(H)$.

Example 3.1.29. Suppose that $H = L^2(T, \mathcal{M}, \mu)$, where $\mathcal{M}$ is countably generated and μ is sigma-finite. Observe that we have a Hilbert space isomorphism

$$H^{\otimes m} \cong L^2\big(T^m, \mathcal{M}^{\otimes m}, \mu^{\otimes m}\big), \quad f_1 \otimes \cdots \otimes f_m \mapsto f_1 \boxtimes \cdots \boxtimes f_m$$

$$f_1 \boxtimes \cdots \boxtimes f_m(t_1, \dots, t_m) = f_1(t_1)\cdots f_m(t_m).$$

Denote by $\mathcal{M}^{\circledcirc n} \subset \mathcal{M}^{\otimes n}$ the σ-algebra generated by the $\mathfrak{S}_n$-invariant measurable functions $T^n \to \mathbb{R}$ consisting of $\mathfrak{S}_n$-invariant $\mathcal{M}^{\otimes n}$-measurable subsets. We set

$$\mu^{\circledcirc n} := \frac{1}{n!}\mu^{\otimes n} : \mathcal{M}^{\otimes n} \to [0, \infty].$$

Observe that $L^2(T, \mathcal{M}, \mu)^{\odot n}$ can be identified with the closed subspace

$$L^2(T^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n})_{\mathfrak{S}_n} \subset L^2(T^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n})$$

consisting of symmetric L^2-functions $F : T^n \to \mathbb{R}$.

The orthogonal projection onto $L^2(T^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n})_{\mathfrak{S}_n}$ is the *symmetrization operator*

$$F \mapsto \mathbf{Sym}(F), \quad \mathbf{Sym}(F)(t_1, \dots, t_n) = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n} F(t_{\sigma(1)}, \dots, t_{\sigma(n)}).$$

For any $f_1, \dots, f_n \in L^2\big(T, \mathcal{M}, \mu\big)$ define

$$f_1 \circledcirc \cdots \circledcirc f_n : T^n \to \mathbb{R},$$

$$\begin{aligned}(f_1 \circledcirc \cdots \circledcirc f_n)(t_1, \dots, t_n) &:= n!\,\mathbf{Sym}(f_1 \boxtimes \cdots \boxtimes f_n)(t_1, \dots, t_n) \\ &= \sum_{\varphi\in\mathfrak{S}_n}\prod_{k=1}^n f_k(t_{\varphi(k)}). = \sum_{\sigma\in\mathfrak{S}_n}\prod_{k=1}^n f_{\sigma(k)}(t_k).\end{aligned} \tag{3.1.38}$$

Clearly, $f_1 \circledcirc \cdots \circledcirc f_n$ is $\mathfrak{S}_n$-invariant.

Lemma 3.1.30. *If $\{f_1, \dots, f_n\} \subset L^2(T, \mu)$ is an orthonormal system and $\alpha \in \mathbb{N}^n$, then*

$$\| \underbrace{f_1 \circledcirc \cdots \circledcirc f_1}_{\alpha_1} \circledcirc \cdots \circledcirc \underbrace{f_n \circledcirc \cdots \circledcirc f_n}_{\alpha_n} \|^2_{L^2(M^{|\alpha|}, \mu^{\circledcirc |\alpha|})} = \alpha!.$$

Proof. We argue as in the proof of Lemma 3.1.27. For any coloring $\pi : \mathbb{I}_{|\alpha|} \to \mathbb{I}_n$ of type α we set

$$F_\pi(t_1, \dots, t_{|\alpha|}) := \prod_{k=1}^{|\alpha|} f_{\pi(k)}(t_k)$$

We have

$$\underbrace{f_1 \circledcirc \cdots \circledcirc f_1}_{\alpha_1} \circledcirc \cdots \circledcirc \underbrace{f_n \circledcirc \cdots \circledcirc f_n}_{\alpha_n} = \alpha! \sum_{\pi\in\mathcal{P}_\alpha} F_\pi$$

Observe that the collection $(F_\pi)_{\pi\in\mathcal{P}_\alpha}$ is orthogonal, and

$$\|F_\pi\|_{L^2(M^{|\alpha|},\mu^{\circledcirc|\alpha|})}=\frac{1}{|\alpha|!},\quad \forall \pi\in\mathcal{P}_\alpha.$$

Hence

$$\|\underbrace{f_1\circledcirc\cdots\circledcirc f_1}_{\alpha_1}\circledcirc\cdots\circledcirc\underbrace{f_n\circledcirc\cdots\circledcirc f_n}_{\alpha_n}\|^2_{L^2(M^{|\alpha|},\mu^{\circledcirc|\alpha|})}=(\alpha!)^2\frac{1}{|\alpha|!}\#\mathcal{P}_\alpha=\alpha!.$$

□

We thus obtain a Hilbert space isomorphism

$$\Psi_n: L^2\big(T,\mathcal{M},\mu\big)^{\hat{\odot}n}\to L^2\big(T^n,\mathcal{M}^{\circledcirc n},\mu^{\circledcirc n}\big) \tag{3.1.39}$$

uniquely determined by the requirement

$$\Psi_n(f_1\odot\cdots\odot f_n)=f_1\circledcirc\cdots\circledcirc f_n,\quad \forall f_1,\dots,f_n\in L^2\big(T,\mathcal{M},\mu\big). \tag{3.1.40}$$

Note that

$$\frac{1}{n!}\int_{T^n}Fd\mu^{\otimes n}=\frac{1}{n!}\int_{T^n}\mathbf{Sym}(F)d\mu^{\otimes n}=\int_{T^n}\mathbf{Sym}(F)d\mu^{\circledcirc n}.$$

□

Remark 3.1.31. Let I denote the unit interval, $\mathcal{B}$ σ-algebra of the Borel subset of I and $\boldsymbol{\lambda}$ the Lebesgue measure on $\mathcal{B}$. For any positive integer n we denote by Δ_n the simplex

$$\Delta_n=\big\{\,(t_1,\dots,t_n)\in I^n;\;\; x_1\leq x_2\leq\cdots\leq x_n\,\big\}.$$

The map

$$L^2\big(I^m,\mathcal{B}_{I^n},\frac{1}{n!}\boldsymbol{\lambda}^{\otimes n}\big)\ni f\to f\big|_{\Delta_n}\in L^2\big(\Delta_m,\mathcal{B}_{\Delta_n}\boldsymbol{\lambda}^{\otimes n}\big)$$

induces a isometric linear isomorphism from the subspace of symmetric functions

$$L^2\big(I^n,\mathcal{B}^{\otimes n},\frac{1}{n!}\boldsymbol{\lambda}^{\otimes n}\big)_{\mathfrak{S}_n}.$$

to $L^2\big(\Delta_n,\mathcal{B}_{\Delta_n},\boldsymbol{\lambda}^{\otimes n}\big)$. □

Suppose that $\mathfrak{X}\subset L^2(\Omega,\mathcal{F},\mathbb{P})$ is a separable Gaussian Hilbert space. We have a linear map

$$\Theta_n:\mathfrak{X}^{\hat{\odot}n}\to\mathfrak{X}^{:n:}, \tag{3.1.41}$$

naturally determined by the correspondences

$$\xi_1\odot\cdots\odot\xi_n\mapsto\xi_1\bullet\cdots\bullet\xi_n=:\xi_1\cdots\xi_n:$$

Corollary 3.1.20, (3.1.17) and (3.1.36) imply that if $\xi_1,\dots,\xi_n$ is an orthonormal system in $\mathfrak{X}$, then

$$\sqrt{n!}\|\;\mathbf{Sym}\big[\,\xi_1\otimes\cdots\otimes\xi_n\big]\,\|=\big\|\,\xi_1\odot\cdots\odot\xi_n\big\|=\big\|:\xi_1\cdots\xi_n:\;\big\|. \tag{3.1.42}$$

We obtain isometries $\Theta_n:\mathfrak{X}^{\odot n}\to\mathfrak{X}^{:n:}$ and thus an isomorphism of graded Hilbert spaces

$$\Theta:\boldsymbol{F}(\mathfrak{X})\to\mathcal{F}(\mathfrak{X}). \tag{3.1.43}$$

The associativity (3.1.31) shows that Θ is actually an isomorphism of algebras. We have thus proved the following result.

Proposition 3.1.32. *The Fock space of a separable Gaussian spaces $\mathfrak{X}$ equipped with the $\odot$ product is isomorphic to the Wiener chaos of $\mathfrak{X}$ equipped with the Wick product $\bullet$.* □

If $\mathfrak{X}_1, \mathfrak{X}_2 \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ are two Gaussian Hilbert spaces, then any bounded linear operator $A : \mathfrak{X}_1 \to \mathfrak{X}_2$ induces bounded linear operators

$$A^{\odot n} : \mathfrak{X}_1^{\hat{\odot} n} \to \mathfrak{X}_2^{\hat{\odot} n}, \;\; n \in \mathbb{N}_0$$

uniquely determined by the requirements

$$A^{\odot n}(\, \xi_1 \odot \cdots \odot \xi_n \,) = (\, A\xi_1 \,) \odot \cdots \odot (\, A\xi_n \,), \;\; \forall \xi_1, \dots, \xi_n \in \mathfrak{X}_1.$$

In particular $A^{\odot 0} = \mathbb{1}$. Moreover

$$\|A^{\odot n}\| = \|A\|^n.$$

If $\|A\| \leq 1$, the operators $A^{\odot n}$ combine to a bounded linear operator

$$\boldsymbol{F}(A) : \boldsymbol{F}(\mathfrak{X}_1) \to \boldsymbol{F}(\mathfrak{X}_2).$$

We deduce that if $\|A\| \leq 1$, then A induces a bounded linear operator $\widehat{A} : \mathcal{F}(\mathfrak{X}_1) \to \mathcal{F}(\mathfrak{X}_2)$ uniquely defined by the equalities

$$\widehat{A}(\, \xi_1 \bullet \cdots \bullet \xi_n \,) = (\, \widehat{A}\xi_1 \,) \bullet \cdots \bullet (\, \widehat{A}\xi_n \,), \;\; \forall \xi_1, \dots, \xi_n \in \mathfrak{X}_1.$$

Note that

$$\|\widehat{A}\| = 1,$$

and we have a commutatiove diagram

$$\begin{array}{ccc} \boldsymbol{F}(\mathfrak{X}_1) & \xrightarrow{\boldsymbol{F}(A)} & \boldsymbol{F}(\mathfrak{X}_2) \\ \Theta \downarrow & & \downarrow \Theta \\ \mathcal{F}(\mathfrak{X}_1) & \xrightarrow[\widehat{A}]{} & \mathcal{F}(\mathfrak{X}_2) \end{array}$$

In particular, a unitary isomorphism $T : \mathfrak{X}_1 \to \mathfrak{X}_1$ induces a canonical unitary isomorphism

$$\widehat{T} : \mathcal{F}(\mathfrak{X}_1) \to \mathcal{F}(\mathfrak{X}_1)$$

which preserves the Wick algebra structure.

Example 3.1.33. Consider the one-dimensional Gaussian Hilbert space $\mathfrak{X}$ spanned by a standard normal random variable ξ. In this case

$$\mathcal{F}(\mathfrak{X}) = L^2(\mathbb{R}, \boldsymbol{\Gamma}).$$

Any linear operator $\mathfrak{X} \to \mathfrak{X}$ has the form $r\mathbb{1}$, and it is a contraction provided $|r| \leq 1$.

Any $f \in L^2(\mathbb{R}, \boldsymbol{\Gamma})$ has the form

$$f(\xi) = \sum_{n \geq 0} f_n H_n(\xi), \;\; f_n = \frac{1}{n!}\mathbb{E}\big[\, f(\xi) H_n(\xi) \,\big] = \frac{1}{n!}\int_{\mathbb{R}} f(x) H_n(x) \boldsymbol{\Gamma}(dx).$$

Since $:\xi^n: := H_n(\xi)$ we deduce that $\widehat{r\mathbb{1}}H_n(\xi) = r^n H_n(\xi)$ and

$$\widehat{r\mathbb{1}}f = \widehat{r\mathbb{1}}\left(\sum_{n\geq 0} f_n H_n(\xi)\right) = \sum_{n\geq 0} f_n r^n H_n(\xi).$$

The operator $\widehat{r\mathbb{1}} : L^2(\mathbb{R}, \mathbf{\Gamma}) \to L^2(\mathbb{R}, \mathbf{\Gamma})$, $|r| \leq 1$ is called the *Mehler transform*. It is an integral operator with kernel

$$\mathcal{M}_r(x, y) = \sum_{n\geq 0} H_n(x)H_n(y)\frac{r^n}{n!} \in L^2(\mathbb{R}^2, \mathbf{\Gamma}^{\otimes 2}).$$

The above series converges uniformly for (x, y, r) on the compacts of $\mathbb{R}^2 \times (-1, 1)$. We denote by $\mathcal{M}_r$ the integral operator $\widehat{r\mathbb{1}}$. Consider the function

$$g_\lambda(x) = \sum_{n\geq 0} H_n(x)\frac{\lambda^n}{n!} = e^{\lambda x - \frac{\lambda^2}{2}}.$$

Observe that for $|r| \leq 1$ we have $\mathcal{M}_r g_\lambda = g_{r\lambda}$. This equality determines $\mathcal{M}_r(x, y)$ uniquely. Consider the function

$$M_r(x, y) = \frac{1}{\sqrt{1 - r^2}} \exp\left(-\frac{(rx)^2 - 2rxy + (ry)^2}{2(1 - r^2)}\right).$$

A direct but tedious computation shows that[5]

$$\int_{\mathbb{R}} M_r(x, y)g_\lambda(y)d\gamma_1(dy) = g_{r\lambda}(x)$$

so that

$$\sum_{n\geq 0} H_n(x)H_n(y)\frac{r^n}{n!} = \frac{1}{\sqrt{1 - r^2}} \exp\left(-\frac{(rx)^2 - 2rxy + (ry)^2}{2(1 - r^2)}\right), \quad \forall |r| < 1.$$

The function $\mathcal{M}_r(x, y) = M_r(x, y)$ is called the *Mehler kernel*.

The family of operators $T_t := \widehat{e^{-t}\mathbb{1}}$, $t \geq 0$ is called the *Ornstein-Uhlenbeck semigroup*. □

3.1.7. Gaussian white noise and the multiple Wiener-Itô integral. Suppose for ease of presentation that $(T, \mathcal{M}, \mu)$ is a *convenient*[6] *probability space*, i.e.,

- The sigma-algebra $\mathcal{M}$ is countably generated, and
- The probability measure μ is complete and non-atomic.

We recall that an *atom* of a measure μ is a measurable set A such $\mu[A] > 0$ and, for any measurable set $B \subset A$, either $\mu[B] = 0$, or $\mu[B] = \mu[A]$. A measure is called *non-atomic* if it has no atoms.

Any non-atomic measure μ enjoys the following property, [**22**, Cor.1.2.10]: for any measurable set A such that $\mu[A] > 0$, and for and $0 < c < \mu[A]$, there exists a measurable set $B \subset A$ such that $\mu[B] = c$.

Let us point out that if $(T, \mathcal{M}, \mu)$ is convenient, then $L^1(T, \mathcal{M}, \mu)$ is separable. Convenient spaces are rather special.

[5] For a more conceptual approach we refer to [**76**, Example 4.18], [**94**, V.1.5] .

[6] The convenient spaces appear in literature with different monikers: standard, perfect, Lebesgue-Rokhlin.

Theorem 3.1.34. *Suppose that* $(T, \mathcal{M}, \mu)$ *is a convenient probability space. Then there exists a measurable function*

$$\Phi : (T, \mathcal{M}) \to \big([0,1], \mathcal{B}_{[0,1]} \big),$$

with the following property.

(i) $\Phi_{\#}\mu = \boldsymbol{\lambda} =$ *the Lebesgue measure on* $[0,1]$.

(ii) *The* μ*-completion of* $\Phi^{-1}\big(\mathcal{B}_{[0,1]}\big)$ *is equal to the* μ*-completion of* $\mathcal{M}$.

□

For proofs and more details we refer to [**23**, Thm. 9.3.4] or [**94**, Thm. IV.4.6.2].

Definition 3.1.35. A (real) *Gaussian white noise* driven by the convenient probability space $(T, \mathcal{M}, \mu)$ is a centered Gaussian process W parametrized by $\mathcal{M}$

$$\Omega \times \mathcal{M} \ni (\omega, A) \mapsto W_\omega\big[\, A \,\big] \in \mathbb{R}$$

with covariance kernel

$$\mathcal{K}_W\big(\, A, B \,\big) = \mathbb{E}\big[\, W\big[\, A \,\big] \cdot W\big[\, B \,\big] \,\big] = \int_T \boldsymbol{I}_A \boldsymbol{I}_B d\mu = \mu\big[\, A \cap B \,\big], \;\; \forall A, B \in \mathcal{M}. \tag{3.1.44}$$

Above $\big(\Omega, \mathcal{S}, \mathbb{P}\big)$ is a probability space. □

Proposition 1.2.10 guarantees the existence of Gaussian white noises.

Fix a Gaussian white noise W driven by $(T, \mathcal{M}, \mu)$. Observe that if $A_1, \dots A_n \in \mathcal{M}$ are pairwise disjoint, then the Gaussian random variable $W\big[\, A_1 \,\big], \dots, W\big[\, A_n \,\big]$ are independent since they are jointly Gaussian and uncorrelated. Observe next that

$$\forall A, B \in M, \;\; A \cap B = \emptyset : \;\; W\big[\, A \,\big] + W\big[\, B \,\big] = W\big[\, A \cup B \,\big] \;\text{ a.s.}. \tag{3.1.45}$$

Indeed

$$\mathbb{E}\Big[\, \big(W\big[\, A \cup B \,\big] - W\big[\, A \,\big] - W\big[\, B \,\big] \big)^2 \,\Big] = \mathbb{E}\Big[\, W\big[\, A \cup B \,\big]^2 \,\Big] + \mathbb{E}\Big[\, W\big[\, A \,\big]^2 \,\Big] + \mathbb{E}\Big[\, W\big[\, B \,\big]^2 \,\Big]$$

$$-2\mathbb{E}\big[\, W\big[\, A \cup B \,\big] W\big[\, A \,\big] \,\big] - 2\mathbb{E}\big[\, W\big[\, A \cup B \,\big] W\big[\, B \,\big] \,\big]$$

$$\stackrel{(3.1.44)}{=} \mu\big[\, A \cup B \,\big] + \mu\big[\, A \,\big] + \mu\big[\, B \,\big] - 2\big(\, \mu\big[\, A \,\big] + \mu\big[\, B \,\big] \,\big) = 0.$$

More generally, if $(A_n)_{n\mathbb{N}}$ is a sequence of pairwise disjoint sets in $\mathcal{M}$ and

$$A = \bigcup_{n \in \mathbb{N}} A_n,$$

then

$$W\big[\, A \,\big] = \sum_{n \in \mathbb{N}} W\big[\, A_n \,\big]$$

where the above convergence is L^2. Kolmogorov's one-series theorem implies that the series also converges a.s..

Remark 3.1.36. One might be tempted to think that the random function $\big(W[A]\big)_{A\in\mathcal{M}}$ admits a modification that is a *random measure*, i.e., for any $\omega\in\Omega$ the correspondence

$$\mathcal{M}\ni A\mapsto W_\omega[A]\in\mathbb{R}$$

is a signed measure. This is not the case. The correct way to view W is as a *stochastic measure*, i.e., *a measure valued in the Hilbert space of random variables* $L^2\big(\Omega,\mathcal{S},\mathbb{P}\big)$. We refer to [85] for a very detailed discussion of the distinction between these two concepts. □

Denote by $\operatorname{Elem}\big(T,\mathcal{M}\big)$ the vector space of $\mathcal{M}$-measurable elementary functions on T, i.e.,

$$\operatorname{Elem}\big(T,\mathcal{M}\big)=\operatorname{span}_{\mathbb{R}}\big\{\boldsymbol{I}_A;\;\; A\in\mathcal{M}\big\}.$$

Using the equality (3.1.45) we deduce that

$$\sum_j a_j\boldsymbol{I}_{A_j}=\sum_k b_k\boldsymbol{I}_{B_k}\Rightarrow\sum_j a_jW[A_j]=\sum_k b_kW[B_k].$$

We thus have a well defined map

$$\mathcal{I}_W:\operatorname{Elem}\big(T,\mathcal{M}\big)\to L^2\big(\Omega,\mathcal{S},\mathbb{P}\big),$$

$$\operatorname{Elem}\big(T,\mathcal{M}\big)\ni f=\sum_j a_j\boldsymbol{I}_{A_j}\mapsto\mathcal{I}_W[f]=\sum_j a_jW[A_j]\in L^2\big(\Omega,\mathcal{S},\mathbb{P}\big).$$

Moreover,

$$\mathbb{E}\big[\mathcal{I}_W[f]^2\big]=\int_T f^2d\mu. \tag{3.1.46}$$

Since the centered Gaussian random variables $W[A_j]$ are independent the random variable $\mathcal{I}_W[f]$ is centered Gaussian.

Since $\operatorname{Elem}\big(T,\mathcal{M}\big)$ is dense in $L^2\big(T,\mathcal{M},\mu\big)$ we deduce from (3.1.46) that $\mathcal{I}_W$ extends to an isometry

$$\mathcal{I}_W:L^2\big(T,\mathcal{M},\mu\big)\to L^2\big(\Omega,\mathcal{S},\mathbb{P}\big)$$

This isometry is called the *Wiener integral* with respect to the white noise W. Traditionally one uses the alternate notation

$$\int_T f(t)W[dt]:=\mathcal{I}_W[f].$$

Since $\mathcal{I}_W$ is an isometry, its image $\mathfrak{X}_W$ is a closed subspace of $L^2\big(\Omega,\mathcal{S},\mathbb{P}\big)$. Each $X\in\mathfrak{X}_W$ is a centered Gaussian random variable since it is a limit of centered Gaussian variables of the form $\mathcal{I}_W[f]$, $f\in\operatorname{Elem}\big(T,\mathcal{M}\big)$. Hence $\mathfrak{X}_W$ is a Gaussian Hilbert space called the *Gaussian Hilbert space associated to the white noise W.*

Conversely, each separable Gaussian Hilbert space is associated to a Gaussian white noise on $(T,\mathcal{M},\mu)$. Indeed, let $\mathfrak{X}\subset L^2\big(\Omega,\mathcal{S},\mathbb{P}\big)$ be a separable Gaussian Hilbert space. The Hilbert space $L^2(T,\mathcal{M},\mu)$ is separable so there exists a (non-canonical) Hilbert space isomorphism

$$\boldsymbol{J}:L^2\big(T,\mathcal{M},\mu\big)\to\mathfrak{X}.$$

Then the map

$$W:\mathcal{M}\to L^2\big(\Omega,\mathcal{S},\mathbb{P}\big),\;\;\mathcal{M}\in A\mapsto W[A]=\boldsymbol{J}[\boldsymbol{I}_A]\in L^2\big(\Omega,\mathcal{S},\mathbb{P}\big)$$

is a Gaussian white noise on $(T,\mathcal{M},\mu)$, and $\boldsymbol{J}=\mathcal{I}_W$.

Suppose that $\mathfrak{X}$ is a separable Gaussian Hilbert space and $\boldsymbol{J} : L^2(T,\mathcal{M},\mu) \to \mathfrak{X}$ is a Hilbert space isomorphism with associated Gaussian white noise W.

Recall that we have associative and commutative products

$$\mathfrak{X}^{:m:} \times \mathfrak{X}^{:n:} \ni (\xi,\eta) \mapsto \xi \bullet \eta := \mathrm{Proj}_{m+n}(\xi\eta) \in \mathfrak{X}^{:(m+n):},$$

$$\odot : \mathfrak{X}^{\odot m} \times \mathfrak{X}^{\odot n} \to \mathfrak{X}^{\odot(m+n)},\ \ X \odot Y := \sqrt{\binom{n+m}{m}}\,\mathbf{Sym}\big(\, X \otimes Y\,\big),$$

The isometries (3.1.41) yield isometries

$$\Theta_n : L^2(T,\mathcal{M},\mu)^{\hat{\odot} n} \to \mathfrak{X}^{:n:},\ \ \forall n \in \mathbb{N}_0.$$

Using the isometries Ψ_n defined in (3.1.39) of Example 3.1.29 we obtain *isometries*

$$J_n : L^2\big(\, T^n, \mathcal{M}^{\circledcirc n}, \mu^{\circledcirc n}\,\big) \xrightarrow{\Psi_n^{-1}} L^2(T,\mathcal{M},\mu)^{\hat{\odot} n} \xrightarrow{\Theta_n} \mathfrak{X}^{:n:}. \tag{3.1.47}$$

Conceptually, J_n maps symmetric functions on T^n to elements in the n-th chaos of $\mathfrak{X} \cong L^2\big(\,\Omega,\mathcal{M},\mu\,\big)$,

J_n : symmetric functions on $T^n \to$ degree n elements in the Fock space $\to$ n-th chaos.

For example, if

$$F = \mathbf{Sym}(f_1 \boxtimes \cdots \boxtimes f_n)(t_1,\dots,t_n) \overset{(3.1.38)}{=} \frac{1}{n!} f_1 \circledcirc \cdots \circledcirc f_n(t_1,\dots,t_n),$$

then

$$\Psi_n^{-1}(F) \overset{(3.1.40)}{=} \frac{1}{n!} f_1 \odot \cdots \odot f_n,$$

and

$$J_n\big[\, F\,\big] = \frac{1}{n!}\boldsymbol{J}\big[\, f_1\,\big] \bullet \cdots \bullet \boldsymbol{J}\big[\, f_n\,\big] = \frac{1}{n!}\big(\ :\boldsymbol{J}\big[\, f_1\,\big]\cdots \boldsymbol{J}\big[\, f_n\,\big]:\ \big).$$

For $F \in L^2(T^n,\mathcal{M}^{\otimes n},\mu^{\otimes n})$ we define its *multiple Wiener-Itô integral* of F to be the *random variable*

$$\mathfrak{I}_n\big[\, F\,\big] := n! J_n\big[\ \mathbf{Sym}(F)\,\big].$$

Often one uses the integral notation

$$\int_{T^n} F dW^n = \int_{T^n} F(t_1,\dots,t_n) W\big[\, dt_1\,\big]\cdots W\big[\, dt_n\,\big] := \mathfrak{I}_n\big[\, F\,\big],\ \ \forall F : M^n \to \mathbb{R}.$$

In particular, if $F \in L^2(T^n,\mathcal{M}^{\otimes n})$ is *symmetric*, then

$$\boxed{\mathfrak{I}_n\big[\, F\,\big] = n! J_n\big[\, F\,\big] \Longleftrightarrow J_n\big[\, F\,\big] = \frac{1}{n!}\int_{T^n} F dW^n.}$$

Note that if $f_1,\dots,f_n \in L^2(M,\mathcal{M},\mu)$, then we obtain the important equality

$$\mathfrak{I}_n\big[\, f_1 \boxtimes \cdots \boxtimes f_n\,\big] = \mathfrak{I}_n\big[\,\mathbf{Sym}\,\big(\, f_1 \boxtimes \cdots \boxtimes f_n\,\big)\,\big] = :\boldsymbol{J}\big[\, f_1\,\big]\cdots \boldsymbol{J}\big[\, f_n\,\big]:. \tag{3.1.48}$$

This equality uniquely determines the multiple Wiener-Itô integral.

Note that since J_n is an isometry we deduce that for any $F \in L^2(M^n,\mathcal{M}^n,\mu^{\otimes n})$ have

$$\mathbb{E}\Big[\,\big|\,\mathfrak{I}_n\big[\, F\,\big]\,\big|^2\,\Big] = \mathbb{E}\Big[\,\big|\,\mathfrak{I}_n\big[\,\mathbf{Sym}\,\big(\, F\,\big)\,\big]\,\big|^2\,\Big] = \mathbb{E}\Big[\,\big|\, n! J_n\big[\,\mathbf{Sym}\,\big(\, F\,\big)\,\big]\,\big|^2\,\Big]$$

$$= \big\|\, n!\,\mathbf{Sym}\,\big(\, F\,\big)\,\big\|^2_{L^2(M^n,\mu^{\circledcirc n})} = n!\|\,\mathbf{Sym}[F]\|^2_{L^2(M^n,\mu^{\otimes n})} \le n!\|F\|^2_{L^2(M^n,\mu^{\otimes n})}.$$

We observe that any $X \in \mathcal{F}(\mathfrak{X})$ has a unique orthogonal decomposition

$$X = \sum_{n\geq 0} \mathfrak{I}_n\big[\, F_n \,\big] = \sum_{n\geq 0} \int_{M^n} F_n dW^n,$$

where $F_n : T^n \to \mathbb{R}$ are symmetric L^2-functions. Moreover

$$\mathbb{E}\big[\, X^2 \,\big] = \sum_{n\geq 0} n! \|\, F_n \,\|^2_{L^2(M^n,\mu^{\otimes n})} = \sum_{n\geq 0} (n!)^2 \| F_n \|^2_{L^2(M^n,\mu^{\odot n})}.$$

Remark 3.1.37. There are many normalization conditions involved in the definition of the multiple Itô integrals and there is danger of confusion since different authors use different conventions.

In [**126**], the Hermite polynomials have a normalization different from the one we use in this book which is the more commonly used.

If $F : T^n \to \mathbb{R}$ is a symmetric function, then $I_n(F)$, as defined in [**76**], [**126**] or [**128**] coincides with the multiple integral $\mathfrak{I}_n\big[\, F \,\big]$ defined above.

The operator J_n that we have described in this section coincides with the operator $\hat{I}_n$ defined in [**76**], the operator $\boldsymbol{I}_G$ defined in [**93**], or the operator J_n in [**128**].. □

Example 3.1.38. (a) Suppose that $f, g \in L^2\big(T, \mathcal{M}, \mu\big)$. Then

$$\int_{T^2} f(x)g(y) W\big[\, dx \,\big] W\big[\, dy \,\big] = \mathfrak{I}_2\big[\, f \boxtimes g \,\big] =: \mathfrak{I}_1\big[\, f \,\big]\mathfrak{I}_1\big[\, g \,\big].$$

We set $X := \mathfrak{I}_1\big[\, f \,\big]$, $Y := \mathfrak{I}_1\big[\, g \,\big]$ so that $X, Y \in \mathfrak{X}$. Then using Wick's formula (3.1.26) we deduce

$$\int_{T^2} f(x)g(y) W\big[\, dx \,\big] W\big[\, dy \,\big] = XY - \mathbb{E}\big[\, XY \,\big]$$

$$= \left(\int_T f(x) W\big[\, dx \,\big] \right) \left(\int_T g(y) W\big[\, dy \,\big] \right) - \int_T f(t)g(t)\mu\big[\, dt \,\big]$$

Thus, the stochastic Fubini formula gets a correction term.

(b) Suppose that $f, g \in L^2\big(T, \mathcal{M}, \mu\big)$. Set $X = \mathfrak{I}_1\big[\, f \,\big]$, $Y = \mathfrak{I}_1\big[\, g \,\big]$. Assume for simplicity that $\|X\|_{L^2} = \|Y\|_{L^2} = 1$.

$$\int_{T^4} f(x_1)f(x_2)g(y_1)g(y_2) W\big[\, dx_1 \,\big] W\big[\, dx_2 \,\big] W\big[\, dy_1 \,\big] W\big[\, dx_2 \,\big]$$

$$= \big(\, : (X \bullet X) \bullet (Y \bullet Y) : \,\big) = (: H_2(X)H_2(Y) :) = (: X^2 Y^2 :)$$

Consider the diagram with vertices

$$X, X, Y, Y$$

If $c = \mathbb{E}\big[\, XY \,\big]$, then using Wick's formula (3.1.26) we deduce

$$: X^2 Y^2 := X^2 Y^2 - \big(\, X^2 + Y^2 + 4cXY \,\big) + 2c^2 + 1.$$

Observe that if $c = 0$, then

$$X^2 Y^2 - \big(\, X^2 + Y^2 + 4cXY \,\big) + 2c^2 + 1 = H_2(X)H_2(Y) = H_{2,2}(X, Y).$$

(c) Suppose that $f_1, \dots, f_k \in L^2\big(T, \mathcal{M}, \mu\big)$ is an *orthonormal* family. Set $X_j := \mathfrak{I}_1\big[\, f_j \,\big]$. Then

$$\mathfrak{I}_{n_1+\cdots+n_k}\big[\, \underbrace{(f_1 \boxtimes \cdots \boxtimes f_1)}_{n_1} \boxtimes \cdots \boxtimes \underbrace{(f_k \boxtimes \cdots \boxtimes f_k)}_{n_k} \,\big] = H_{n_1}\big(X_1\big) \cdots H_{n_k}(X_k). \tag{3.1.49}$$

How does one go about computing the multiple Wiener-Itô integrals in general? Theoretically one goes as follows. Fix complete orthonormal system $(f_n)_{n\in\mathbb{N}}$ of $L^2\big(T, \mathcal{M}, \mu\big)$. Set $X_n = \mathfrak{I}_1\big[\, f_n \,\big]$.Then the collection

$$\underbrace{(f_{m_1} \boxtimes \cdots \boxtimes f_{m_1})}_{n_1} \boxtimes \cdots \boxtimes \underbrace{(f_{m_k} \boxtimes \cdots \boxtimes f_{m_k})}_{n_k}$$

$k \in \mathbb{N}$, $m_1 < \cdots < m_k$, $n_1, \cdots, n_k \in \mathbb{N}$, $n_1 + \cdots + n_k = n$ is a complete orthonormal system of $L^2\big(T^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n}\big)$ and, using (3.1.49) one can theoretically comute $\mathfrak{I}_n\big[\, F \,\big]$ for any $F \in L^2\big(T^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n}\big)$. For this to work in concrete situations we need to be very lucky. When $L^2\big(T, \mathcal{M}, \mu\big) = L^2\big([0,1], \mathcal{B}_{[0,1]}, \boldsymbol{\lambda}\big)$ and the white noise isomorphism $\mathfrak{I}_1$ is given by the Brownian white noise

$$f \mapsto \mathfrak{I}_1\big(\mathfrak{I}_1\big[\, f \,\big] = \int_0^1 f(t)dB(t),$$

then one can use Itô's stochastic calculus to compute multiple Wiener-Itô integrals; see Remark 3.1.43. □

Suppose that $\mathfrak{X}$ is a separable Gaussian Hilbert space and $\boldsymbol{J} : L^2(T, \mathcal{M}, \mu) \to \mathfrak{X}$ is a Hilbert space isomorphism with associated Gaussian white noise W.

If $F : T^2 \to \mathbb{R}$ is an integrable function we define the *contraction*

$$CF := \int_T F(t,t)\mu\big[\, dt \,\big].$$

More generally, if $F : T^n \to \mathbb{R}$ and $1 \leq i < j \leq n$ we define the *contraction* $C_{ij}F : T^{n-2} \to \mathbb{R}$ to be

$$C_{ij}F := \int_T F(t_1, \dots, t_n)_{t_i = t_j = t}\, \mu\big[\, dt \,\big].$$

Given a Feynman diagram $\Gamma \in \mathrm{Feyn}(n)$ we set

$$C_\Gamma := \prod_{\boldsymbol{e}\in\mathcal{E}(\Gamma)} C_{\boldsymbol{e}},$$

where for any edge $\boldsymbol{e} = (i,j)$ of Γ we set $C_{\boldsymbol{e}} = C_{ij}$.

Lemma 3.1.39. *Suppose we are given functions $F_i \in L^2(T^{n_i})$, $i = 1, \dots, k$. We define*

$$F = F_1 \boxtimes \cdots \boxtimes F_k : T^n \to \mathbb{R}, \;\; n = \sum_{i=1}^k n_i,$$

$$F_1 \boxtimes \cdots \otimes F_k(t_{11}, \dots, t_{1n_1}; \dots; t_{k1}, \dots, t_{kn_k}) := \prod_{j=1}^k F_j\big(t_{j1}, \dots, t_{jn_k}\big).$$

For any Feynman diagram Γ compatible with the array

$$(i,j), \;\; 1 \leq i \leq k, \;\; 1 \leq j \leq n_i,$$

we have

$$\| C_\Gamma(F_1 \boxtimes \cdots \boxtimes F_k) \|_{L^2(M^{n-2r(\Gamma)})} \leq \prod_{j=1}^k \| F_j \|_{L^2}.$$

Proof. We use induction on k. The case $k = 1$ is trivial since Γ has no wedge and thus $C_\Gamma(F) = F$.

For $k = 2$, we can assume, after relabeling the variables, that the $r(\Gamma)$ edges of Γ connectd the vertices $(1, j)$ and $(2, j)$, $j = 1, \dots, r$. Then for

$$t' \in T^{n_1 - r}, \;\; t'' \in T^{n_2 - r}, \;\; s \in T^r$$

we have

$$C_\Gamma F(t', t'') = \int_{T^r} F_1(s, t') F_2(s, t'') \mu^r \big[\, ds \,\big]$$

and thus, by Cauchy-Schwarz

$$|C_\Gamma F(t', t'')|^2 \leq \left(\int_{T^r} F_1(s, x')^2 \mu^r \big[\, ds \,\big] \right) \left(\int_{T^r} F_2(s, t'')^2 \right) \mu^r \big[\, ds \,\big]$$

Integrating the remaining variables (t', t'') we deduce

$$\| C_\Gamma(F) \|_{L^2}^2 \leq \| F_1 \|_{L^2}^2 \cdot \| F_2 \|_{L^2}^2.$$

This disposes of the case $k = 2$.

For $k > 2$ we set

$$F_2' = f_2 \boxtimes \cdots \boxtimes F_k.$$

Denote by Γ_1 the subdiagram of Γ consisting of the edges that have one vertex on the first row, $(1, j)$, $1 \leq j \leq n_1$, and denote by Γ_2 the subdiagram of Γ determined by the edges of Γ that connect points on rows different from the first row.

We than have

$$C_\Gamma(F) = C_{\Gamma_1}\big(F_1 \boxtimes C_{\Gamma_2}(F_2') \big).$$

Thus, using the inequality established for $k = 2$ and the induction assumption we reach the desired conclusion. □

Suppose we are given functions $F_i \in L^2(T^{n_i})$, $i = 1, \dots, k$. We set

$$Y_i = \mathfrak{I}_{n_i}\big[\, F_i \,\big] = \mathfrak{I}_{n_i}\big[\, \mathbf{Sym}(F_i) \,\big] \in \mathfrak{X}^{:n_i:}, \;\; 1 \leq i \leq k.$$

From the Diagram Formula (Theorem 3.1.23) we deduce via a simple density argument (see [**76**, Thm. 7.33] for details) the following important result.

Theorem 3.1.40. *We set $n := n_1 + \cdots + n_k$. Denote by $\mathcal{A}$ the array*

$$(i, j), \;\; 1 \leq i \leq k, \;\; 1 \leq j \leq n_i.$$

Then

$$Y_1 \cdots Y_k = \sum_{\Gamma \in \mathrm{Feyn}\,[\,\mathcal{A}\,]} \mathfrak{I}_{n-2r(\Gamma)}\big[\, C_\Gamma(F_1 \boxtimes \cdots \boxtimes F_k) \,\big], \tag{3.1.50}$$

In particular

$$\mathbb{E}\big[\, Y_1 \cdots Y_k \,\big] = \sum_{\Gamma \in \mathrm{Feyn}^*\,[\,\mathcal{A}\,]} C_\Gamma(F_1 \boxtimes \cdots \boxtimes F_k). \tag{3.1.51}$$

Proof. Lemma 3.1.39 shows that all the contractions and stochastic integrals are well defined. From the definition of the multiple Wiener-Itô integrals we deduce that the right-hand side of (3.1.50) defines a continuous multilinear map

$$\prod_{i=1}^{k} L^2(T^{n_i}, \mu^{n_i}) \to \mathcal{F}(\mathfrak{X}).$$

Thus it suffice to verify the equality (3.1.50) in the special case when each f_i is a monomial

$$f_i(t_1, \dots, t_{m_i}) = f_{i1}(t_1) \cdots f_{im_i}(t_{m_i}).$$

This special case follows immediately from the diagram formula (3.1.28a). □

Remark 3.1.41. Theorem 3.1.40 corresponds to [**93**, Thm. 5.3] where it is referred to as the *Diagram Formula.* □

Suppose that $F_i \in L^2\big(T^{n_i}\big)$, $i = 1, 2$, are symmetric functions and $r \leq \min(n_1, n_2)$. We define $F_1 \boxtimes_r F_2 : T^{n_1+n_2-2r} \to \mathbb{R}$

$$F_1 \boxtimes_r F_2(t_1, \dots, t_{n_1-r}, s_1, \dots, s_{n_2-r}) := \int_{T^r} F_1\big(\, t_1, \dots t_{n_1-r}, \boldsymbol{t}\,\big) F_2(\boldsymbol{t}, s_1, \dots s_{n_2-r}\,) \mu^r\big[\, d\boldsymbol{t}\,\big].$$

The equality (3.1.50) in the case $k = 2$ can now be rewritten as

$$\mathfrak{I}_{n_1}\big[\, F_1\,\big]\mathfrak{I}_{n_2}\big[\, F_2\,\big] = \sum_{r=0}^{\min(n_1,n_2)} r! \binom{n_1}{r} \binom{n_2}{r} \mathfrak{I}_{n_1+n_2-2r}\big[\, F_1 \boxtimes_r F_2\,\big]. \tag{3.1.52}$$

Digression 3.1.42. The multiple Wiener-Itô integral can be given a constructive description that justifies the terminology *integral.* I will describe below the contours of this construction. For details refer to [**95**, §VI.2] or [**126**, §1.1.2].

Suppose that $(T, \mathcal{M}, \mu)$ is a convenient probability space and W is a Gaussian white noise driven by $(T, \mathcal{M}, \mu)$. Denote by $\mathfrak{X}$ the Gaussian space generated by W. Assume that $\mathcal{M}$ is generated by the measurable sets M_n, $n \in \mathbb{N}$. I secretly think that $T = [01,]$ and $\mathcal{M} = \mathcal{B}_{[0,1]}$.

For $N \in \mathbb{N}$ we denote by $\mathcal{M}_N$ the sigma-subalgebra of $\mathcal{M}$ generated by $M_1, \dots, M_N$. We set $\mathcal{M}_0 := \{\emptyset, T\}$. For any $N \geq 0$, the sigma-subalgebra $\mathcal{M}_N$ consists of finitely many measurable subsets of T and hence it corresponds to the sigma-algebra generated by a finite, measurable partition of T. I secretly think that $\mathcal{M}_N$ is the sigma algebra, generated by the partition

$$\big[\,(k-1)2^{-N}, k2^{-N}\,\big), \;\; k = 1, \dots, 2^N - 1, \;\; \big[\,1 - 2^{-N}, 1\,\big].$$

The parts of this partition are precisely the atoms of $\mu|_{\mathcal{M}_N}$. Moreover,

$$\mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \;\; \text{and} \;\; \mathcal{M} = \bigvee_{N \geq 0} \mathcal{M}_N.$$

Denote by $\boldsymbol{A}_N$ the set of atoms of $\mu|_{\mathcal{M}_N}$. Fix $p \in \mathbb{N}$ and set

$$\mathcal{M}_N^p := \mathcal{M}_N^{\otimes p}.$$

Each of the sigma-subalgebras $\mathcal{M}_N^p$ is finite and its set of atoms is

$$\boldsymbol{A}_{N,p} := \underbrace{\boldsymbol{A}_N \times \cdots \times \boldsymbol{A}_N}_{p}.$$

Set

$$H_{N,p} := L^2\big(T^p, \mathcal{M}_N^p, \mu^{\otimes p}\big),\ \ H_p = L^2\big(T^p, \mathcal{M}^{\otimes p}, \mu^{\otimes p}\big).$$

Any function $F \in H_{N_np}$ is constant on the set of atoms $\boldsymbol{A}_{N,p}$ and we denote by $F(\underline{\boldsymbol{a}})$ its value on the atom $\underline{\boldsymbol{a}} \in \boldsymbol{A}_{N,p}$. Then

$$F = \sum_{\underline{\boldsymbol{a}} \in \boldsymbol{A}_{N,p}} F(\underline{\boldsymbol{a}}) \boldsymbol{I}_{\underline{\boldsymbol{a}}}$$

and

$$\|F\|_{L^2}^2 = \sum_{\underline{\boldsymbol{a}} \in \boldsymbol{A}_{N,p}} F(\underline{\boldsymbol{a}})^2 \mu^{\otimes p}\big[\,\underline{\boldsymbol{a}}\,\big].$$

The symmetric group $\mathfrak{S}_p$ acts in the obvious way on the set of atoms $\underline{\boldsymbol{a}} = \boldsymbol{a}_1 \times \cdots \times \boldsymbol{a}_p \in \boldsymbol{A}_{N,p}$. More precisely, for $\sigma \in \mathfrak{S}_p$.

$$\sigma \cdot \underline{\boldsymbol{a}} = \boldsymbol{a}_{\sigma(1)} \times \cdots \times \boldsymbol{a}_{\sigma(p)}.$$

For $\underline{\boldsymbol{a}} \in \boldsymbol{A}_{N,p}$ we denote by $\operatorname{Stab}(\underline{\boldsymbol{a}})$ the stabilizer of $\underline{\boldsymbol{a}}$ with respect to the action of $\mathfrak{S}_p$. More explicitly,

$$\operatorname{Stab}(\underline{\boldsymbol{a}}) := \big\{\,\sigma \in \mathfrak{S}_p;\ \ \sigma \cdot \underline{\boldsymbol{a}} = \underline{\boldsymbol{a}}\,\big\}.$$

We set

$$\boldsymbol{A}^*_{N,p} := \big\{\,\underline{\boldsymbol{a}} \in \boldsymbol{A}_{N,p};\ \ \operatorname{Stab}(\underline{\boldsymbol{a}}) = \{1\}\,\big\},\ \ \boldsymbol{A}^0_{N,p} := \boldsymbol{A}_{N,p} \setminus \boldsymbol{A}^*_{N,p}.$$

Note that

$$\underline{\boldsymbol{a}} = \boldsymbol{a}_1 \times \cdots \times \boldsymbol{a}_p \in \boldsymbol{A}^0_{N,p} \Longleftrightarrow \exists i \neq j\ \ \boldsymbol{a}_i = \boldsymbol{a}_j.$$

We set

$$X_{N,p} := \big\{\,F \in H_{N,p};\ \ F(\underline{\boldsymbol{a}}) = 0,\ \ \forall \underline{\boldsymbol{a}} \in \boldsymbol{A}^0_{N,p}\,\big\},$$

We have natural inclusions $i_N : H_{N,p} \hookrightarrow H_{N+1,p}$. Note that $i_N\big(X_{N,p}\big) \subset X_{N+1,p}$.

For $F \in X_{N,p}$ we set

$$\mathfrak{I}_{p,N}\big[\,F\,\big] = \sum_{\underline{\boldsymbol{a}} \in \boldsymbol{A}^*_{N,p}} F(\underline{\boldsymbol{a}}) W\big[\,\underline{\boldsymbol{a}}\,\big],\ \ W\big[\,\underline{\boldsymbol{a}}\,\big] := W\big[\,\boldsymbol{a}_1\,\big] \cdots W\big[\,a_p\,\big] \in \mathcal{F}(\mathfrak{X}).$$

The Gaussian random variables $W\big[\,\boldsymbol{a}_1\,\big], \dots, W\big[\,\boldsymbol{a}_p\,\big]$ are independent elements of $\mathfrak{X}$ so that

$$W\big[\,\underline{\boldsymbol{a}}\,\big] \in \mathfrak{X}^{:p:}$$

and thus $\mathfrak{I}_{p,N}\big[\,F\,\big] \in \boldsymbol{X}^{:p:}$. Since $\mathfrak{S}_p$ acts freely on $\boldsymbol{A}^*_{N,p}$ and

$$W\big[\,\sigma \cdot \underline{\boldsymbol{a}}\,\big] = W\big[\,\underline{\boldsymbol{a}}\,\big],\ \ \forall \underline{\boldsymbol{a}} \in \boldsymbol{A}_{N,p},\ \ \forall \sigma \in \mathfrak{S}_p$$

we deduce

$$\mathfrak{I}_{p,N}\big[\,F\,\big] = \mathfrak{I}_{p,N}\big[\,\mathbf{Sym}(F)\,\big] = \sum_{\underline{\boldsymbol{a}} \in \boldsymbol{A}^*_{N,p}} \left(\frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} F(\sigma \cdot \underline{\boldsymbol{a}})\right) W\big[\,\underline{\boldsymbol{a}}\,\big].$$

The operators $\mathfrak{I}_{p,N}$ are compatible with the inclusions $i_N : X_{N,p} \hookrightarrow X_{N+1,p}$, i.e., for any $N \geq 0$, the following diagram is commutative

$$\begin{array}{ccc} X_{N,p} & \xhookrightarrow{\ i_N\ } & X_{N+1,p} \\ & {\scriptstyle \mathfrak{I}_{p,N}} \searrow \quad \swarrow {\scriptstyle \mathfrak{I}_{p,N+1}} & \\ & \mathfrak{X}^{:p:} & \end{array}$$

We have thus constructed a well defined map

$$\mathfrak{I}_p : X_p \to \mathfrak{X}^{:p:},\ \ X_p = \bigcup_{N\geq 0} X_{N,p} \in L^2\big(T^p, \mathcal{M}^{\otimes p}, \mu^{\otimes p}\big).$$

We set $\tilde{F} = \mathbf{Sym}(F)$. Then

$$\mathfrak{I}_p\big[\,\tilde{F}\,\big]\cdot\mathfrak{I}_p\big[\,\tilde{F}\,\big] = \mathfrak{I}_p\big[\,F\,\big]\cdot\mathfrak{I}_p\big[\,F\,\big] = \sum_{\underline{\boldsymbol{a}},\underline{\boldsymbol{a}}'\in\boldsymbol{A}^*_{N,p}} \tilde{F}(\underline{\boldsymbol{a}})\tilde{F}(\underline{\boldsymbol{a}}')W\big[\,\underline{\boldsymbol{a}}\,\big]W\big[\,\underline{\boldsymbol{a}}'\,\big].$$

Now observe that

$$\mathbb{E}\big[\,W\big[\,\underline{\boldsymbol{a}}\,\big]W\big[\,\underline{\boldsymbol{a}}'\,\big]\,\big] = \begin{cases}\mu^{\otimes p}\big[\,\underline{\boldsymbol{a}}\,\big], & \underline{\boldsymbol{a}}'\in\mathfrak{S}_p\cdot\underline{\boldsymbol{a}},\\ 0, & \text{otherwise.}\end{cases}$$

We deduce

$$\big\|\,\mathfrak{I}_p\big[\,\tilde{F}\,\big]\,\big\|^2 = p!\sum_{\boldsymbol{a}\in\boldsymbol{A}^*_{N,p}}\tilde{F}(\boldsymbol{a})^2\mu^{\otimes p}\big[\,\underline{\boldsymbol{a}}\,\big] = p!\big\|\,\tilde{F}\,\big\|^2_{L^2} \leq p!\big\|\,F\,\big\|^2_{L^2}.$$

We have thus produced a bounded linear map

$$\mathfrak{I}_p : X_p \to \mathfrak{X}^{:p:}.$$

Using the fact that μ is non-atomic one can show that X_p is dense in H_p so $\mathfrak{I}_p : X_p \to \mathfrak{X}^{:p:}$ admits a unique extension as a bounded linear map $\mathfrak{I}_p : H_p \to \mathfrak{X}^{:p:}$.

We need to explain why in the above proof we have the avoided the atoms $\underline{\boldsymbol{a}} = \boldsymbol{a}_1\times\cdots\times\boldsymbol{a}_p$ that intersect the diagonal, i.e., $\boldsymbol{a}_i = \boldsymbol{a}_j$ for some $i\neq j$.

To see what goes wrong consider the simplest case $p=2$, $T=[0,1]$ and $\mu=\boldsymbol{\lambda}$. The map $\mathfrak{I}_2$ should be a map

$$\mathfrak{I}_2 : L^2\big(T^2\big)\to\mathfrak{X}^{:2:},\ \ F\mapsto\mathfrak{I}_2\big[\,F\,\big]$$

In particular,

$$\mathbb{E}\big[\,\mathfrak{I}_2\big[\,F\,\big]\,\big] = 0.$$

Clearly, for any non-negligible Borel subset $B\subset T$, we have

$$\mathbb{E}\big[\,W\big[\,B\times B\,\big]\,\big] = \mathrm{Var}\,\big[\,W\big[\,B\,\big]\,\big] = \boldsymbol{\lambda}\big[\,B\,\big]\neq 0.$$

We set

$$\boldsymbol{a}_{k,n} := \begin{cases}[(k-1)2^{-n}, k2^{-n}), & 1\leq k<2^n,\\ \big[\,1-2^{-n},1\,\big], & k=n,\end{cases}$$

and

$$\boldsymbol{A}_n = \big\{\,\boldsymbol{a}_{1,n},\dots,\boldsymbol{a}_{2^n,n}\,\big\},\ \ \mathcal{A}_n = \sigma\big(\,\boldsymbol{A}_n\,\big)$$

For each $F\in L^2(T^2)$ we set $F_n = \mathbb{E}\big[\,F\,\|\,\mathcal{A}_{n,2}\,\big]\in L^2(T^2,\mathcal{A}_{n,2})$. Then

$$F_n = \sum_{\underline{\boldsymbol{a}}\in\boldsymbol{A}_{n,2}} F(\underline{\boldsymbol{a}})\boldsymbol{I}_{\underline{\boldsymbol{a}}}$$

and the Martingale Convergence Theorem shows that F_n converges in L^2 to F. One might be tempted to set

$$\mathfrak{I}_2\big[\,F\,\big] = \lim_{n\to\infty}\sum_{\underline{\boldsymbol{a}}\in\boldsymbol{A}_{n,2}} F(\underline{\boldsymbol{a}})W\big[\,\underline{\boldsymbol{a}}\,\big].$$

Suppose that $F=\boldsymbol{I}_{T^2}$. Note that for any $\boldsymbol{a}\in\boldsymbol{A}_n$ we have

$$\mathbb{E}\big[\,W\big[\,\boldsymbol{a}\times\boldsymbol{a}\,\big]\,\big] = \mathrm{Var}\,\big[\,W\big[\,\boldsymbol{a}\,\big]\,\big] = 2^{-n},\ \ \forall\boldsymbol{a}\in\boldsymbol{A}_n$$

$$\mathrm{Var}\big[\,W\big[\,\boldsymbol{a}\times\boldsymbol{a}\,\big]\,\big]=3\cdot 2^{-2n}-2^{-2n}=2\cdot 2^{-2n}=2^{-2n+1}.$$

In this case If we set

$$S_n=\sum_{\boldsymbol{a}\in\boldsymbol{A}_n}W\big[\,\boldsymbol{a}\times\boldsymbol{a}\,\big],$$

then

$$\mathbb{E}\big[\,S_n=1\,\big]=1,\ \ \mathrm{Var}\big[\,S_n\,\big]=2^{-n+1}.$$

and we deduce from the Borel-Cantelli lemma that S_n converges a.s. to 1 as $n\to\infty$. Note that $S_n\in H_n\setminus X_n$. Thus

$$\mathbb{E}\big[\,\boldsymbol{I}_{T^2}\,\|\,\mathcal{A}_{n,2}\,\big]=\sum_{\underline{\boldsymbol{a}}\in\boldsymbol{A}_{n,2}}W\big[\,\underline{\boldsymbol{a}}\,\big]=S_n+\sum_{\underline{\boldsymbol{a}}\in\boldsymbol{A}^*_{n,2}}W\big[\,\underline{\boldsymbol{a}}\,\big]$$

It is not difficult to see that

$$\sum_{\underline{\boldsymbol{a}}\in\boldsymbol{A}^*_{n,2}}W\big[\,\underline{\boldsymbol{a}}\,\big]=W\big[\,T\,\big]^2-1=H_2\big(W\big[\,T\,\big]\,\big).$$

In general, given $F\in L^2(T^2)$ we set

$$F_n^{\Delta}=\sum_{\boldsymbol{a}\in\boldsymbol{A}_n}F_n\big(\,\boldsymbol{a}\times\boldsymbol{a}\,\big)\boldsymbol{I}_{\boldsymbol{a}\times\boldsymbol{a}},$$

then

$$\lim_{n\to\infty}\mathrm{Var}\,\big[\,F_n^{\Delta}\big]=\int_{T^2}F\Delta_{\#}\boldsymbol{\lambda},$$

where $\Delta:T\to T^2$ is the diagonal map. For a more in-depth discussion of this aspect and generalizations we refer to [**54, 85, 131**]. □

Remark 3.1.43. When $T=[0,1]$, $\mathcal{M}=\mathcal{B}_{[0,1]}$, $\mu=\boldsymbol{\lambda}$ one can give an alternate description of the multiple Wiener-Itô integrals. Fix a Brownian motion $(B(t))_{t\in[0,1]}$ and denote by $\mathfrak{X}$ the Gaussian spaces spanned by the random function $t\mapsto B(t)$. The white noise W is the more familiar white noise $W=dB(t)$ and the white noise isomorphism is given by the Itô integral

$$L^2\big(\,[0,1]\,\big)\ni f\mapsto\mathfrak{I}_1\big(\,F\,\big)=\int_0^1 f(t)dB(t).$$

Note that

$$B(t)=\mathfrak{I}_1\big[\,\boldsymbol{I}_{[0,t]}\,\big].$$

Above, we interpret $f(t)$ as a predictable process in the obvious way. In particular

$$\mathfrak{I}_1\big[\,\boldsymbol{I}_{[t_0,t_1]}\,\big]=B(t_1)-B(t_0).$$

Since the Haar wavelets span a dense subset of $L^2\big(\,[0,1]\,\big)$ we deduce that the above map $\mathfrak{I}_1:L^2\big(\,0,1]\,\big)\to\mathfrak{X}$ is indeed an isomorphism. For $n\geq 1$ we set

$$\Delta_n:=\big\{\,(t_1,\dots,t_n);\ \ t_1\leq t_2\leq\cdots\leq t_n\,\big\}.$$

Any symmetric function $F:[0,1]^{\to}\mathbb{R}$ is uniquely determined by its restriction to Δ_n. One can prove that for any $F\in L^2\big(\,[0,1]^n,\mathcal{B}_{[0,1]^n},\boldsymbol{\lambda}^{\otimes n}\,\big)$ we have

$$\mathfrak{I}_n\big[\,F\,\big]=\mathfrak{I}_n\big[\,\mathbf{Sym}(F)\,\big]=\int_{t_n}^1 dB(s_n)\int_{t_{n-1}}^{t_n}dB(s_{n-1})\cdots\int_0^{t_1}F\big(\,s_1,\dots,s_n\,\big)dB(s_1).\tag{3.1.53}$$

For $n \geq 1$, the right-hand-side of the above equality is an integral Itô integral.

I want to highlight the main ideas of the proof of (3.1.53) in [**76**, Thm.7.5] to which I refer for details.

The proof is inductive. For any $(t_1, \dots, t_{n-1}) \in \Delta_{n-1}$, any $F \in L^2\big(\Delta_n\big)$, and any $t \in [0,1]$ define $F_t : \Delta_{n-1} \to \mathbb{R}$

$$F_t(t_1, \dots, t_{n-1}) = \begin{cases} f(t_1, \dots, t_{n-1}, t), & t \geq t_{n-1}, \\ 0, & t < t_{n-1}. \end{cases}$$

One shows that for any $n \geq 2$

$$\mathfrak{I}_n\big[\,F\,\big] = \int_0^1 \hat{\mathfrak{I}}_{n-1}\big[\,F_t\,\big]dB(t),$$

where the right-hand-side above is the Itô integral of a predictable process. The above equality is linear and continuous in F so it can be reduced to the case when F is the indicator of a box contained in Δ_n. In this case it can be verified by direct computation.

The equalities (3.1.16) and (3.1.53) show that

$$\mathfrak{I}_n\big[\,\boldsymbol{I}_{[0,t]^n}\,\big] = H_n\big(\,B(t)\,\big|t\,\big). \tag{3.1.54}$$

In particular,

$$\mathfrak{I}_n\big[\,\boldsymbol{I}_{[0,1]^n}\,\big] = H_n\big(\,B(1)\,\big).$$

□

3.2. Malliavin calculus

3.2.1. The Malliavin gradient and Gaussian Sobolev spaces. Suppose H is a separable real Hilbert space, and $\mathfrak{X} \subset L^2(\Omega, \mathcal{S}, \mathbb{P})$ is a separable real Gaussian Hilbert space.

- We denote by $L^0_{\mathfrak{X}}(\Omega)$ the space of $\mathcal{S}_{\mathfrak{X}}$-measurable functions $f : \Omega \to \mathbb{R}$ modulo a.s. equality. $(X_n)_{n\in\mathbb{N}}$ is a dense subset of $\mathfrak{X}$, then for any $f \in L^0_{\mathfrak{X}}(\Omega)$ there exists a measurable function $F : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ such that
$$f(\omega) = F\big(\,X_1(\omega), \dots, X_n(\omega), \dots\,\big), \;\; \forall \omega \in \Omega.$$
- For $p \in [1, \infty]$ we denote by $L^p_{\mathfrak{X}}(\Omega)$ the subspace of $L^0_{\mathfrak{X}}(\Omega)$ consisting of p-integrable functions equipped with the usual L^p-norm. Note that $\mathcal{F}(\mathfrak{X}) = L^2_{\mathfrak{X}}(\Omega)$. If
- We denote by $L^0_{\mathfrak{X}}(\Omega, H)$ the space of $\mathcal{S}_{\mathfrak{X}}$-measurable maps $f : \Omega \to H$ modulo a.s. equality, and by $L^p_{\mathfrak{X}}(\Omega, H)$ the subspace $L^0_{\mathfrak{X}}(\Omega, H)$ consisting of maps $f : \Omega \to H$ such that $\|f\| \in L^p_{\mathfrak{X}}(\Omega)$. The norm in this space is
$$\mathbb{E}\big[\,\|f\|^p_H\,\big]^{\frac{1}{p}}.$$
The space $L^0_{\mathfrak{X}}(\Omega, H)$ is equipped with a bilinear map
$$\begin{gathered} [-,-]_H : L^0_{\mathfrak{X}}(\Omega, H) \times L^0_{\mathfrak{X}}(\Omega, H) \to L^0_{\mathfrak{X}}(\Omega), \\ [f,g]_H(\omega) := \big(\,f(\omega), g(\omega)\,\big)_H, \;\; \forall f, g \in L^0_{\mathfrak{X}}(\Omega, H). \end{gathered} \tag{3.2.1}$$
- A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *admissible* if it is smooth, and its derivatives, of any order, have at most polynomial growth.

- An H-valued function $f:\mathbb{R}^n\to$ is called *admissible* if it is smooth, and its derivatives, of any order, have at most polynomial growth.

The random variables $F\in L^0_{\mathfrak{X}}\big(\,\Omega\,\big)$ are commonly referred as (nonlinear) functionals of the Gaussian space $\mathfrak{X}$. They all can be *non-uniquely* expressed as

$$F=\varphi\big(\,X_1,\dots,X_n,\dots\,\big),$$

where $\varphi:\mathbb{R}^{\mathbb{N}}\to\mathbb{R}$ is measurable and $X_n\in\mathfrak{X}$, $\forall n$.

We will construct various Banach subspaces $L^0_{\mathfrak{X}}(\Omega)$. These depend only on $\mathfrak{X}$.

Definition 3.2.1. Let $\mathfrak{X}\subset L^2\big(\,\Omega,\mathcal{S},\mathbb{P}\,\big)$ be a separable real Gaussian Hilbert space.

(i) We denote by $\mathfrak{A}(\mathfrak{X})\subset L^0_{\mathfrak{X}}(\Omega)$ the set of random variables of the form $f(X_1,\dots,X_m)$, where $m\in\mathbb{N}$, $f:\mathbb{R}^m\to\mathbb{R}$ is an admissible function, and $X_1,\dots,X_m\in\mathfrak{X}$.

(ii) We denote by $\mathcal{P}(\mathfrak{X})\subset\mathfrak{A}(\mathfrak{X})$ the set of random variables of the form $P(X_1,\dots,X_m)$, where $P:\mathbb{R}^m\to\mathbb{R}$ is a polynomial in m variables with real coefficients and $X_1,\dots,X_m\in\mathfrak{X}$.

(iii) Suppose that H is a separable Hilbert space. We denote by $\mathfrak{A}(\mathfrak{X},H)$ the subspace of $L^0_{\mathfrak{X}}\big(\,\Omega,H\,\big)$ spanned by random vectors of the form $f(X_1,\dots,X_n)$, where $f:\mathbb{R}^n\to H$ is admissible and $X_1,\dots,X_N\in\mathfrak{X}$.

□

Note that if $(X_n)_{n\geq}$ is a complete orthonormal system of $\mathfrak{X}$, then for any $k\in\mathbb{N}$ and any $n_1,\dots,n_k\in\mathbb{N}_0$ the polynomial $H_{n_1}(X_1)\cdots H_{n_k}(X_k)$ belongs to $L^p_{\mathfrak{X}}\big(\,\Omega\,\big)$, $\forall p\in[1,\infty)$. In particular,

$$\mathcal{P}(\mathfrak{X}),\mathfrak{A}(\mathfrak{X})\in L^p_{\mathfrak{X}}\big(\,\Omega\,\big),\;\;\forall p\in[1,\infty).$$

Arguing exactly as in the proof of Proposition 3.1.15 we deduce the following result

Proposition 3.2.2. *The spaces $\mathcal{P}(\mathfrak{X})$ and $\mathfrak{A}(\mathfrak{X})$ are dense in $L^p_{\mathfrak{X}}\big(\,\Omega\,\big)$ for any $p\in[1,\infty)$.* □

For $X\in\mathfrak{X}$ and $f(X_1,\dots,X_m)\in\mathfrak{A}(\mathfrak{X})$ we define $D_Xf(X_1,\dots,X_m)\in L^0_{\mathfrak{X}}(\Omega)$ by setting

$$D_Xf(X_1,\dots,X_m)(\omega):=\sum_j\frac{\partial f}{\partial x_j}\big(\,X_1(\omega),\dots,X_m(\omega)\,\big)(X_j,X)_{\mathfrak{X}},\tag{3.2.2}$$

where $(-,-)_{\mathfrak{X}}$ denotes the inner product in $\mathfrak{X}$, $(X,Y)_{\mathfrak{X}}=\mathbb{E}[\,XY\,]$. We have the a.s. equality

$$D_Xf:=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big(\,f\big(\,X_1+\varepsilon(X_1,X)_{\mathfrak{X}},\dots,X_m+\varepsilon(X_m,X)_{\mathfrak{X}}\,\big)-f\big(\,X_1,\dots,X_m\,\big)\Big).\tag{3.2.3}$$

From the definition (3.2.2) it is not clear whether the equality

$$f(X_1,\dots,X_m)=g(Y_1,\dots,Y_n)\in\mathfrak{A}(\mathfrak{X}),$$

where f and g are admissible, implies that

$$D_Xf(X_1,\dots,X_m)=D_Xg(Y_1,\dots,Y_n)\in L^0_{\mathfrak{X}}(\Omega),\;\;\forall X\in\mathfrak{X}.$$

This is indeed the case, but the proof is more involved. The key fact is that the a.s. equality

$$f(X_1,\dots,X_m)=g(Y_1,\dots,Y_n)\in\mathfrak{A}(\mathfrak{X})$$

implies that, for any $X \in \mathfrak{X}$, we have

$$f\big(X_1 + (X, X_1)_{\mathfrak{X}}, \dots, X_m + (X, X_m)_{\mathfrak{X}}\big) = g\big(Y_1 + (X, Y_1)_{\mathfrak{X}}, \dots, Y_n + (X, Y_n)_{\mathfrak{X}}\big).$$

For details we refer to [**76**, Thm.14.1& Def.15.26]. We will have more to say about this in Digression 3.2.9.

Given $f(X_1, \dots, X_n) \in \mathfrak{A}(\mathfrak{X})$, define $Df(X_1, \dots, X_m) \in L^0_{\mathfrak{X}}(\Omega, \mathfrak{X})$

$$Df(X_1, \dots, X_m)(\omega) = \sum_j \frac{\partial f}{\partial x_j}\big(X_1(\omega), \dots, X_m(\omega)\big) X_j(-). \tag{3.2.4}$$

Equivalently, Df is the unique $\mathcal{S}_{\mathfrak{X}}$-measurable map $T : \Omega \to \mathfrak{X}$ such that

$$\big[T, X\big]_{\mathfrak{X}} = D_X f \;\text{ a.s.}, \quad \forall X \in \mathfrak{X},$$

where $\big[-,-\big]_{\mathfrak{X}}$ is the bilinear map defined in (3.2.1). The resulting operator

$$f(X_1, \dots, X_m) \mapsto Df(X_1, \dots, X_m)$$

is called the *Malliavin gradient* or *derivative*.

Example 3.2.3. Let $X \in \mathfrak{X}$. Then DX is the constant map $\Omega \to \mathfrak{X}$, $\omega \mapsto X$. For this reason we will rewrite (3.2.4) in the form

$$Df(X_1, \dots, X_m)(\omega) = \sum_j \frac{\partial f}{\partial x_j}\big(X_1(\omega), \dots, X_m(\omega)\big) DX_j. \tag{3.2.5}$$

This notation better conveys the nature of the two factors $\frac{\partial f}{\partial x_j}\big(X_1(\omega), \dots, X_m(\omega)\big)$ and DX_j. The first is a *scalar*, while the second is an element of $\mathfrak{X}$. Note also that

$$D_X F = \big[DF, DX\big]_{\mathfrak{X}} \in L^0_{\mathfrak{X}}(\Omega)$$

where $[-,-]_{\mathfrak{X}}$ is the bilinear form (3.2.1). □

For a positive integer k and $f(X_1, \dots, X_m)$ as above we define

$$D^k f(X_1, \dots, X_m) \in L^0_{\mathfrak{X}}\big(\Omega, \mathfrak{X}^{\hat{\odot} p}\big)$$

by setting

$$D^k f(X_1, \dots, X_m)(\omega) = \sum_{i_1, \dots, i_k = 1}^{m} \frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_k}}\big(X_1(\omega), \dots, X_m(\omega)\big) DX_{i_1} \otimes \cdots \otimes DX_{i_k}.$$

Remark 3.2.4. Arguing as in Lemma 3.1.27 we deduce that

$$\sum_{i_1, \dots, i_k = 1}^{m} \frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_k}}\big(X_1(\omega), \dots, X_m(\omega)\big) DX_{i_1} \otimes \cdots \otimes DX_{i_k}$$

$$= \sqrt{k!} \sum_{\substack{\alpha \in \mathbb{N}_0^m, \\ |\alpha| = k}} \frac{1}{\alpha!} \partial_x^\alpha f\big(X_1(\omega), \dots, X_m(\omega)\big) (DX_1)^{\odot \alpha_1} \odot \cdots \odot (DX_m)^{\odot \alpha_m}$$

$$= \sum_{\substack{\alpha \in \mathbb{N}_0^m, \\ |\alpha| = k}} \frac{1}{\alpha!} \partial_x^\alpha f\big(X_1(\omega), \dots, X_m(\omega)\big) \, \mathbf{Sym}\big[(DX_1)^{\otimes \alpha_1} \otimes \cdots \otimes (DX_m)^{\otimes \alpha_m}\big]. \quad □$$

Observe that the class $\mathfrak{A}(\mathfrak{X})$ contains the algebra generated by the polynomials $H_n(X)$, $X \in \mathfrak{X}$. Arguing as in the proof of Proposition 3.1.15 we deduce that $\mathfrak{A}(\mathfrak{X})$ is dense in $L^q_{\mathfrak{X}}(\Omega, \mathbb{P})$, $\forall q \in (1, \infty)$.

Proposition 3.2.5. *Let $k \in \mathbb{N}$ and $q \in (1, \infty)$. Then the operator*

$$D^k : \mathfrak{A}(\mathfrak{X}) \subset L^q_{\mathfrak{X}}(\Omega) \to L^q_{\mathfrak{X}}\big(\Omega, \mathfrak{X}^{\hat{\odot} p}\big)$$

is closable.

Proof. We follow closely the proof of [**124**, Prop.2.3.4]. We consider only the case $k = 1$.

Let $F, G \in \mathfrak{A}(\mathfrak{X})$ and $X \in \mathfrak{X}$ such that $\|X\|_{L^2} = 1$. Note that $FG \in \mathfrak{A}(X)$. We can assume that

$$F = f(X_1, \dots, X_n), \;\; G = g(X_1, \dots, X_n).$$

where $\{X_1, \dots, X_n\} \subset \mathfrak{X}$ is an orthonormal system, $X_1 = X$, and f, g are admissible. Then, setting $h = fg$ we deduce

$$\mathbb{E}\big[\,\big[\,D(FG), X\,\big]_{\mathfrak{X}}\,\big] = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \frac{\partial h}{\partial x_1}(x_1, \dots, x_n) e^{-\frac{x_1^2 + \cdots + x_n^2}{2}} dx$$

(integrate by parts along the x_1-direction)

$$= (2\pi)^{-n/2} \int_{\mathbb{R}^n} x_1 h(x_1, \dots, x_n) e^{-\frac{x_1^2 + \cdots + x_n^2}{2}} dx = \mathbb{E}\big[\,XFG\,\big].$$

Clearly $D(FG) = G(DF) + F(DG)$. We deduce the following Gaussian integration by parts formula

$$\mathbb{E}\big[\,G\big[\,DF, X\,\big]_{\mathfrak{X}}\,\big] = -\mathbb{E}\big[\,F\big[\,DG, X\,\big]_{\mathfrak{X}}\,\big] + \mathbb{E}\big[\,XFG\,\big], \;\; \forall X \in \mathfrak{X}. \tag{3.2.6}$$

Using the notation (3.2.2) we can rewrite the above equality in the more suggestive form

$$\mathbb{E}\big[\,(D_X F)G\,\big] = \mathbb{E}\big[\,F(-D_X + X)G\,\big], \;\; \forall X \in \mathfrak{X}. \tag{3.2.7}$$

The above equation extends by linearity to all $X \in \mathfrak{X}$, not necessarily of L^2-norm 1.

Now let (F_n) be a sequence in $\mathfrak{A}(\mathfrak{X})$ such that the following hold.

(i) $F_n \to 0$ in $L^q_{\mathfrak{X}}(\Omega)$.

(ii) The sequence DF_n converges in the norm of $L^q_{\mathfrak{X}}(\Omega, \mathfrak{X})$ to some $\eta \in L^q_{\mathfrak{X}}(\Omega, \mathfrak{X})$.

We have to show that $\eta = 0$ a.s.. Let $X \in \mathfrak{X}$, $G \in \mathfrak{A}(\mathfrak{X})$. Since $F_n \to 0$ in L^q and XG and $\big[\,DG, X\,\big]_{\mathfrak{X}}$ belong to $L^{\frac{q}{q-1}}$ we deduce from (3.2.6) that

$$\mathbb{E}\big[\,G\big[\,\eta, X\,\big]_{\mathfrak{X}}\,\big] = \lim_{n\to\infty} \mathbb{E}\big[\,G\big[\,DF_n, X\,\big]_{\mathfrak{X}}\,\big]$$

$$= -\lim_{n\to\infty} \mathbb{E}\big[\,F_n\big[\,DG, X\,\big]_{\mathfrak{X}}\,\big] + \lim_{n\to\infty} \mathbb{E}\big[\,XF_nG\,\big] = 0.$$

Thus

$$\mathbb{E}\big[\,G\big[\,\eta, X\,\big]_{\mathfrak{X}}\,\big] = 0, \;\; \forall G \in \mathfrak{A}(\mathfrak{X}), \;\; X \in \mathfrak{X}.$$

Since $\mathfrak{A}(\mathfrak{X})$ is dense in any L^r, $r \in [1, \infty)$, we deduce that

$$\forall X \in \mathfrak{X}, \;\; (\eta, X)_{\mathfrak{X}} = 0 \;\text{ a.s.}.$$

Thus, if $(e_k)_{k\in\mathbb{N}}$ is an orthonormal basis of $\mathfrak{X}$, there exists a negligible set $\mathcal{N} \subset \Omega$ such that

$$\big[\,\eta, e_n\,\big]_{\mathfrak{X}}(\omega) = 0, \;\; \forall n \in \mathbb{N}, \;\; \omega \in \Omega \setminus \mathcal{N}.$$

Thus $\eta = 0$ a.s. □

Definition 3.2.6. Let $k \in \mathbb{N}$ and $q \in [1,\infty)$. We define the *Gaussian Sobolev space* $\mathbb{D}^{k,q}(\mathfrak{X})$ to be the closure of $\mathfrak{A}(\mathfrak{X})$ with respect to the norm

$$\|F\|_{\mathbb{D}^{k,q}} := \left(\sum_{j=0}^{k} \mathbb{E}\Big[\, \|D^j F\|^q_{\mathfrak{X}^{\hat{\odot} k}} \, \Big] \right)^{\frac{1}{q}}. \qquad \square$$

Remark 3.2.7. According to Proposition 3.2.5, the operator D^k can be consistently extended as a continuous operator

$$D^k : \mathbb{D}^{k,q}(\mathfrak{X}) \to L^q_{\mathfrak{X}}(\Omega, \mathfrak{X}^{\hat{\odot} p}).$$

The space $\mathbb{D}^{k,q}(\mathfrak{X})$ is the domain of the closure of the unbounded operator

$$D^k : \mathfrak{A}(\mathfrak{X}) \subset L^q_{\mathfrak{X}}(\Omega) \to L^q_{\mathfrak{X}}(\Omega, \mathfrak{X}^{\hat{\odot} p}),$$

i.e., the closure of $\mathfrak{A}(\mathfrak{X})$ in the graph norm of D^k. In particular, the space $\mathfrak{A}(\mathfrak{X})$ is dense in $\mathbb{D}^{k,q}(\mathfrak{X})$, $\forall k \geq 0$, $q \in [1,\infty)$.

The space $\mathbb{D}^{k,2}(\mathfrak{X})$ is a Hilbert space with inner product

$$(F,G)_{\mathbb{D}^{p,2}} = \sum_{j=0}^{k} \mathbb{E}\big[\, \big[\, D^j F, D^j G \, \big]_{\mathfrak{X}^{\hat{\odot} j}} \, \big].$$

$\square$

Remark 3.2.8. Let $(\Omega, \mathcal{S}, \mathbb{P})$ be a probability space. Suppose that $(\boldsymbol{T}, \mathcal{M}, \mu)$ is a convenient probability space and

$$W : H := L^2(\boldsymbol{T}, \mathcal{M}, \mu) \to L^2\big(\Omega, \mathcal{S}, \underline{\ }\big)$$

is a Gaussian white noise. Its image is a Gaussian Hilbert space $\mathfrak{X}$. If $F \in \mathbb{D}^{1,2}(\mathfrak{X})$ then its gradient is an element in

$$L^2\big(\Omega, \mathcal{S}_{\mathfrak{X}}, \mathbb{P}; \mathfrak{X}\big) \cong L^2\big(\Omega, \mathcal{S}_{\mathfrak{X}}, \mathbb{P}; H\big) \cong L^2\big(\Omega \times \boldsymbol{T}, \mathbb{P} \otimes \mu\big),$$

and thus it can be identified with a stochastic process parametrized by $\boldsymbol{T}$. We denote by $D_t F$ this stochastic process.

Any functional F in the n-th Wiener chaos $\mathfrak{X}^{:n:}$ can be written as multiple Wiener integral

$$F = \mathcal{I}_n\big[\, f_n \,\big] = \int_{\boldsymbol{T}^n} f_n(t_1, \dots, t_n) W\big[\, dt_1 \,\big] \cdots W\big[\, dt_n \,\big],$$

where $f_n \in L^2\big(\boldsymbol{T}, \mathcal{M}^{\otimes n}, \mu^{\otimes n}\big)$. The gradient DF can be identified with the stochastic process

$$D_t F = n \mathcal{I}_{n-1}\big[\, f_n(-,t) \,\big] = n \int_{\boldsymbol{T}^{n-1}} f_n(t_1, \dots, t_{n-1}, t) W\big[\, dt_1 \,\big] \cdots W\big[\, dt_{n-1} \,\big].$$

I refer to [**126**, Sec 1.2.1] for a proof and more information on this point of view.

For example if $F = \mathcal{I}_n\big[\, \boldsymbol{I}_{[0,1]^n} \,\big] = H_n\big(B(1)\big)$, then the equality (3.1.54) shows that

$$D_t F = n H_{n-1}\big(B(t) \,\big|\, t\big).$$

One can show that for any $f \in L^2\big([0,1]\big)$ and any $n \in \mathbb{N}$ we have (see [**128**, Sec.3.2.1])

$$D_t \left(\int_0^1 f(s) dB(s) \right)^n = \left(\int_0^1 f(s) dB(s) \right)^{n-1} f(t).$$

Thus, if we set $X = \mathcal{I}_1\big[\, f \,\big]$, then $DX = D_t \mathcal{I}_1\big[\, f \,\big] = f(t)$.

For a more suggestive description of D_tF I refer to [**96**, Sec. 1.2-1.3]. □

Digression 3.2.9. The usual Sobolev spaces can be defined in two equivalent, yet qualitatively different ways: as completions of spaces of smooth functions with respect to Sobolev or by directly describing the regularity conditions that characterize the functions belonging to a given Sobolev space.

Similarly, the Gaussian Sobolev spaces $\mathbb{D}^{k,p}(\mathfrak{X})$ can be given an alternate definition by describing explicitly the regularity properties a random variable in $L^0_{\mathfrak{X}}(\Omega)$ needs to satisfy in order to belong to $\mathbb{D}^{k,p}$. I digress to offer the reader an idea of this approach. I follow closely [**76**, Chap.15] to which we refer for proofs and more details.

A random variable $Z \in L^0_{\mathfrak{X}}(\Omega)$ can be described non-uniquely as

$$Z(\omega) = f\big(X_1(\omega), \dots, X_n(\omega), \dots\big)$$

where $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ is a measurable function and $X_n \in \mathfrak{X}$, $\forall n$. Suppose we are given another such description of Z

$$Z = g\big(Y_1, \dots, Y_n, \dots\big)$$

with $g : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ a measurable function and $Y_n \in \mathfrak{X}$, $\forall n$. One can show that for any $X \in \mathfrak{X}$ we have

$$f\big(X_1 + \mathbb{E}[X_1X], \dots, X_n + \mathbb{E}[X_nX], \dots\big) = g\big(Y_1 + \mathbb{E}[Y_1X], \dots, Y_n + \mathbb{E}[Y_nX], \dots\big).$$

The key fact behind this equality is the identity

$$\mathbb{E}\big[\varphi\big(X_1 + \mathbb{E}[X_1X], \dots, X_n + \mathbb{E}[X_nX]\big)\big] = \mathbb{E}\big[:e^X: \varphi\big(X_1, \dots, X_n\big)\big]$$

for any $n \in \mathbb{N}$ and any bounded measurable function $\varphi : \mathbb{R}^n \to \mathbb{R}$. Above, $:e^X:$ denotes the Wick exponential defined in Example 3.1.25 . We then have a well defined linear map

$$\rho_X : L^0_{\mathfrak{X}}(\Omega) \to L^0_{\mathfrak{X}}(\Omega)$$

given by

$$\rho_X f\big(X_1, \dots, X_n, \dots\big) = f\big(X_1 + \mathbb{E}[X_1X], \dots, X_n + \mathbb{E}[X_nX], \dots\big).$$

This is also a morphism of algebras, i.e.,

$$\rho_X\big(Z_1Z_2\big) = \rho_X\big(Z_1\big)\rho_X\big(Z_2\big), \;\; \forall X \in \mathfrak{X}, \;\; \forall Z_1, Z_2 \in L^0_{\mathfrak{X}}(\Omega).$$

Moreover, $\rho_{X_1+X_2} = \rho_{X_1} \circ \rho_{X_2}$, $\forall X_1, X_2 \in \mathfrak{X}$. This map is called the *Cameron-Martin shift*. It satisfies many other pleasant properties [**76**, Thm. 14.1].

Given $F \in L^0_{\mathfrak{X}}(\Omega)$ and $X \in \mathfrak{X}$ we say that the directional derivative $\partial_X F$ exists if the difference quotients

$$\frac{1}{t}\big(\rho_{tX}F - F\big)$$

converge in probability as $t \to 0$. The limit is the directional derivative $\partial_X F$. If $f \in C^1\big(\mathbb{R}^n\big)$ $X, X_1, \dots, X_n \in \mathfrak{X}$, then

$$\partial_X f\big(X_1, \dots, X_n\big) = \sum_{k=1}^n \mathbb{E}[X_kX] \frac{\partial f}{\partial x_k}\big(X_1, \dots, X_n\big).$$

A functional $F \in L^0_{\mathfrak{X}}\big(\Omega\big))$ is said to have a *gradient* if there exists $G \in L^0_{\mathfrak{X}}\big(\Omega, \mathfrak{X}\big)$ such that for any $X \in \mathfrak{X}$ the directional derivative $\partial_X F$ exists and

$$\partial_X F = \big[\, G, X \,\big]_{\mathfrak{X}} \text{ a.s..}$$

If this happens we set $DF := G$.

A random variable $F \in L^0_{\mathfrak{X}}\big(\Omega\big)$ is said to be *absolutely continuous along* $X \in X$, or X-a.c. if there exists a version of $t \mapsto \rho_{tX} F$ such that for any $\omega \in \Omega$ the function $t \mapsto \rho_{tX} F(\omega)$ is absolutely continuous. We say that $F \in L^0_{\mathfrak{X}}\big(\Omega\big)$ is *ray absolutely continuous* or *ray a.c.* if it is X-a.c for any $X \in \mathfrak{X}$ We denote by $\tilde{\mathbb{D}}^{1,0}\big(\mathfrak{X}\big)$ the space of functionals $F \in L^0\big(\Omega\big)$ that are ray a.c. and admit a gradient.

Following [**76**, Def. 15.59] we define $\tilde{\mathbb{D}}^{1,p}$ to be the the subspace of $\tilde{\mathbb{D}}^{1,0}\big(\mathfrak{X}\big)$ consisting of functionals F such that $DF \in L^p$. The space $\tilde{\mathbb{D}}^{1,p}\big(\mathfrak{X}\big)$ is equipped with an obvious Sobolev-type L^p-norm. The fact that the normed space $\tilde{\mathbb{D}}^{1,p}\big(\mathfrak{X}\big)$ is isomorphic to the Banach space $\mathbb{D}^{1,p}\big(\mathfrak{X}\big)$ in Definition 3.2.6 requires some work. For details we refer to [**76**, Thm. 15.104].

The space $\tilde{\mathbb{D}}^{1,p}\big(\mathfrak{X}\big)$ has certain technical advantages. In particular, it leads naturally to the following result with important applications.

Theorem 3.2.10. *Suppose that $F \in \mathbb{D}^{1,p}(\mathfrak{X})$, $p \in [1, \infty)$, is a non-constant random variable. Then its distribution is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.* □

For a proof I refer to [**76**, Thm. 15.50]. Yet other approaches to this absolute continuity theorem can be found in [**25**], [**95**, Thm. III.7.1] or [**144**].

Let me mention a few things about a special case of Gaussian Hilbert spaces arising frequently in stochastic analysis. Suppose that $\mathbf{\Gamma}$ is a Gaussian measure on the separable Fréchet space $\boldsymbol{X}$. It has an associated Gaussian Hilbert space $\mathfrak{X}$. More precisely, $\mathfrak{X} = \boldsymbol{X}^*_{\mathbf{\Gamma}}$, the closure of the image of the tautological map

$$T_{\mathbf{\Gamma}} : \boldsymbol{X}^* \to L^2(\boldsymbol{X}, \mathbf{\Gamma})$$

defined in (1.1.28). Concretely $\boldsymbol{X}^*_{\mathbf{\Gamma}}$ can be identified with the quotient of $\boldsymbol{X}^*$ modulo $\mathbf{\Gamma}$-a.s. equality. In this case $\big(\Omega, \mathcal{S}, \mathbb{P}\big) = \big(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}}, \mathbf{\Gamma}\big)$. For $X \in \mathfrak{X}$ the Cameron-Martin shift τ_X coincides with the pullback induced by a measurable map of $\tau_X : \boldsymbol{X} \to \boldsymbol{X}$.

In (1.1.34) we defined Cameron-Martin space $H_{\mathbf{\Gamma}} = T^*_{\mathbf{\Gamma}}\big(\mathfrak{X}\big) \subset \boldsymbol{X}$ of $\mathbf{\Gamma}$. For $X \in \mathfrak{X}$ define

$$\tau_X : \boldsymbol{X} \to \boldsymbol{X}, \;\; \tau_X(x) = x + T^*_{\mathbf{\Gamma}} X.$$

Then, for any $F \in \boldsymbol{X}^*$, the Cameron-Martin shift $\rho_X F$ is given by

$$\rho_X(F) = F + \mathbb{E}\big[\, XF \,\big] \overset{(1.1.29)}{=} F + F\big(T^*_{\Gamma} X\big) = \tau^*_X(F).$$

Recall that F not really a function, but an equivalence class of functions modulo equality Γ-a.s. Thus if $F = F'$ Γ-a.s., but $F \neq F'$, it is possible that $F\big(x + T^*_{\mathbf{\Gamma}} X\big) \neq F'\big(x + T^*_{\mathbf{\Gamma}} X\big)$ $\mathbf{\Gamma}$-a.s.. This not the case.

The classical Cameron-Martin theorem [**32**] shows that the measure $(\tau_X)_{\#}\mathbf{\Gamma}$ is absolutely continuous with respect to $\mathbf{\Gamma}$. More precisely

$$(\tau_X)_{\#}\Gamma\big[\, dx \,\big] = e^{X(x) - \frac{1}{2}\|X\|_{L^2(\boldsymbol{X},\Gamma)}}\Gamma\big[\, dx \,\big].$$

For modern presentations we refer to [**21**, Sec. 2.4] or [**148**, Sec. 3.3.1].

If $\boldsymbol{X} = \mathbb{R}^n$ and $\Gamma = \Gamma_{\mathbb{1}_n}$, then $\mathfrak{X} = \mathbb{R}^n$, and for any $X \in \mathbb{R}^n$, the Cameron-Martin shift $\tau_X : \mathbb{R}^n \to \mathbb{R}^n$ is the translation

$$\mathbb{R}^n \ni \boldsymbol{v} \mapsto \boldsymbol{v} + X \in \mathbb{R}^n.$$

If $\boldsymbol{X} = C\big([0,1]\big)$ and $\boldsymbol{\Gamma}$ is a Gaussian measure on $\boldsymbol{X}$, then the linear functionals

$$\mathbf{Ev}_t : C\big([0,1]\big) \to \mathbb{R}, \;\; \mathbf{Ev}_t(f) = f(t)$$

span a dense subspace of the associated Gaussian Hilbert space $\boldsymbol{X}^*_\Gamma$. The associated Cameron-Martin space is the same of the associated Cameron-Martin space of the continuous Gaussian process $\big(\mathbf{Ev}_t\big)_{t\in[0,1]}$.

For example, if $\boldsymbol{\Gamma}$ is the Wiener measure on $C([0,1])$, then the Gaussian process $(\mathbf{Ev}_t)_{t\in[0,1]}$ is the Brownian motion.

As explained in Example B.5.5, to each $\xi \in \boldsymbol{X}^*$ we can naturally associate a continuous function $h_\xi : [0,1] \to \mathbb{R}$, $h_\xi(t) = \mathbb{E}\big[\,\xi\,\mathbf{Ev}_t\,\big]$, $\forall t$. We obtain a translation

$$\mathcal{T}_\xi : \boldsymbol{X} \to \boldsymbol{X}, \;\; f \mapsto f + h_\xi.$$

We deduce that

$$\mathcal{T}_\xi^*\,\mathbf{Ev}_t(f) = \mathbf{Ev}_t(f) + \mathbf{Ev}_t(h_\xi) = \mathbf{Ev}_t(f) + \mathbb{E}\big[\,\xi\,\mathbf{Ev}_t\,\big]$$

i.e.,

$$\mathcal{T}_\xi^*\,\mathbf{Ev}_t = \mathbf{Ev}_t + \mathbb{E}\big[\,\xi\,\mathbf{Ev}_t\,\big].$$

Since the collections $\mathbf{Ev}_t$ span a dense subspace of $\boldsymbol{X}^*_\Gamma$ we deduce that

$$\forall \eta \in \boldsymbol{X}^*_{\boldsymbol{\Gamma}} \;\; \mathcal{T}_\xi^*\eta = \eta + \mathbb{E}\big[\,\xi\eta\,\big].$$

Thus, the pullback induced by the translation in $\boldsymbol{X}$ by h_ξ is the Cameron-Martin shift ρ_ξ.

This ends the digression. □

Example 3.2.11. Suppose that $\mathfrak{X}$ is a finite dimensional Gaussian Hilbert space, $\dim \mathfrak{X} = n$. Fix an orthonormal basis $X_1, \dots, X_n$. Then

$$L^2_{\mathfrak{X}}(\Omega) \cong L^2\big(\mathbb{R}^n, \boldsymbol{\Gamma}_{\mathbb{1}}[dx]\big), \;\; \boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dx\,\big] = (2\pi)^{-n/2} e^{-\frac{|x|^2}{2}}\,dx.$$

If $f \in C^\infty(\mathbb{R}^n)$ is a function such that is derivatives of any order have at most polynomial growth, then the Malliavin gradient $Df(X_1, \dots, X_n)$ corresponds to the differential of f

$$df = \sum_{k=1}^n \frac{\partial f}{\partial x_k} dx_k.$$

Furthermore, the Gaussian Sobolev space corresponds to the weighted Sobolev space $W^{k,q}\big(\mathbb{R}^n, \boldsymbol{\Gamma}_{\mathbb{1}}\big)$ equipped with the norm

$$\|f\|_{\mathbb{D}^{k,q}} = \left(\sum_{|\alpha|\leq k} \int_{\mathbb{R}^n} |\partial_x^\alpha f(x)|^q\, \boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dx\,\big] \right)^{\frac{1}{q}}.$$ □

Proposition 3.2.12. *Let $f \in \mathcal{F}(\mathfrak{X})$, $p \in \mathbb{N}$. Recall that* Proj_n *denotes the orthogonal projection onto the n-th chaos $\mathfrak{X}^{:n:}$. The following statements are equivalent.*

(i) $F \in \mathbb{D}^{p,2}(\mathfrak{X})$.

(ii)

$$\sum_{n\geq 0} n^p \| \operatorname{Proj}_n F\|^2 < \infty.$$

Outline of proof. Fix an orthonormal basis $\underline{X} = (X_k)_{k\geq 1}$ of $\mathfrak{X}$. We have

$$\operatorname{Proj}_n = \sum_{\alpha\in\mathbb{N}_0,\ |\alpha|=n} c_\alpha(f) H_\alpha(\underline{X}).$$

From the equality (3.1.4b) we deduce that

$$\int_{\mathbb{R}^N} \partial_j H_\alpha(x) \partial_j H_\beta(x) \mathbf{\Gamma}(dx) = \alpha_j \big(H_\alpha, H_\beta \big)_{L^2(\mathbf{\Gamma})}.$$

This implies that

$$\|DH_\alpha\|^2_{L^2} = |\alpha| \|H_\alpha\|_{L^2}.$$

In particular, we deduce that

$$\mathfrak{X}^{:n:} \subset \mathbb{D}^{1,2}(\mathfrak{X}),\ \ \|F\|^2_{\mathbb{D}^{1,2}} = (1+n)\|F\|^2_{L^2},\ \ \forall F \in \mathfrak{X}^{:n:}.$$

The proposition is now an immediate consequence of the above fact. □

Example 3.2.13. For any $n \in \mathbb{N}$, and any $p \in \mathbb{N}_0$, the n-th chaos $\mathfrak{X}^{:n:}$ is contained in $\mathbb{D}^{p,2}(\mathfrak{X})$. □

Since $\mathfrak{A}(\mathfrak{X})$ is dense in in $\mathbb{D}^{1,q}(\mathfrak{X})$ we obtain the following useful result.

Proposition 3.2.14 (Chain Rule)**.** *Suppose that $\varphi : \mathbb{R}^m \to \mathbb{R}$ is a C^1-function with bounded derivatives. Then for any $F_1, \dots, F_n \in \mathbb{D}^{1,q}$ we have $\varphi(F_1, \dots, F_n) \in \mathbb{D}^{1,p}$ and*

$$D\varphi(F_1, \dots, F_m) = \sum_{j=1}^m \frac{\partial \varphi}{\partial x_j}(F_1, \dots, F_m) DF_m. \tag{3.2.8}$$

□

The Chain Rule holds in the more general case when φ is a Lipschitz function, [**126**, Prop. 1.2.4].

Proposition 3.2.15 (Extended Chain Rule)**.** *Suppose that $\varphi : \mathbb{R}^m \to \mathbb{R}$ is a Lipschitz function, then for any $F_1, \dots, F_n \in \mathbb{D}^{1,q}$ such that the probability distribution of*

$$\vec{F} = (F_1, \dots, F_m) : \Omega \to \mathbb{R}^m,$$

is absolutely continuous[7] *with respect to the Lebesgue measure on $\mathbb{R}^m$, then $\varphi(\vec{F}) \in \mathbb{D}^{1,q}$ and (3.2.8) continues to hold with $\frac{\partial\varphi}{\partial x_i}$ defined a.e.* □

[7]This assumption is needed to give a precise meaning to $\frac{\partial\varphi}{\partial x_i}(\vec{F})$ since for a Lipschitz function φ the partial derivatives $\frac{\partial\varphi}{\partial x_i}(x)$ are defined only for x outside a Lebesgue negligible subset $\mathcal{N} \subset \mathbb{R}^m$.

3.2.2. The divergence operator. The *divergence operator* δ is the adjoint of the Malliavin gradient viewed as a closed *unbounded* operator

$$D:\mathbb{D}^{1,2}(\mathfrak{X})\subset L^2_{\mathfrak{X}}(\Omega)\to L^2_{\mathfrak{X}}\big(\Omega,\mathfrak{X}\big).$$

Similarly, for $p\in\mathbb{N}$, the operator δ^p is the adjoint of the closed unbounded operator

$$D^p:\mathbb{D}^{p,2}(\mathfrak{X})\subset L^2_{\mathfrak{X}}(\Omega)\to L^2_{\mathfrak{X}}\big(\Omega,\mathfrak{X}^{\hat{\odot}p}\big).$$

The domain $\operatorname{Dom}(\delta^p)$ of δ^p is the space

$$\Big\{u\in L^2_{\mathfrak{X}}(\Omega,\mathfrak{X}^{\hat{\odot}p});\;\;\exists C>0\;\;\big|\mathbb{E}\big[\,\big[\,D^pF,u\,\big]_{\mathfrak{X}^{\hat{\odot}p}}\,\big]\big|\le C\sqrt{\mathbb{E}\big[\,F^2\,\big]},\;\;\forall F\in\mathfrak{A}(\mathfrak{X})\,\Big\}.$$

If $u\in\operatorname{Dom}(\delta^p)$, then δ^pu is the unique element in $L^2_{\mathfrak{X}}(\Omega)=L^2_{\mathfrak{X}}\big(\,\Omega\,\big)$ such that

$$\mathbb{E}\big[\,F\delta^pu\,\big]=\mathbb{E}\big[\,\big[\,D^pF,u\,\big]_{\mathfrak{X}^{\hat{\odot}p}}\,\big],\;\;\forall F\in\mathfrak{A}(\mathfrak{X}).\tag{3.2.9}$$

Example 3.2.16. (a) Suppose that $\dim\mathfrak{X}=n<\infty$. Fix an orthonormal basis $\{X_1,\dots,X_n\}$ of $\mathfrak{X}$. Let

$$u=(u_1,\dots,u_n)\in L^2_{\mathfrak{X}}\big(\,\Omega,\mathbb{R}^n\,\big).$$

Then each u_j is a measurable function of $(X_1,\dots,X_n)$. For any admissible function $f\in C^\infty(\mathbb{R}^n)$ we have

$$\mathbb{E}\big[\,f(X_1,\dots,X_n)\delta u\,\big]=\mathbb{E}\left[\sum_{j=1}^n f'_{x_j}(X_1,\dots,X_n)u_j(X_1,\dots,X_n)\right]$$

$$=(2\pi)^{-n/2}\int_{\mathbb{R}^n}\left(\sum_{j=1}^n f'_{x_j}(x)u_j(x)e^{-\frac{x_j^2}{2}}\right)dx=\int_{\mathbb{R}^n}f(x)\sum_{j=1}^n\big(-\partial_{x_j}u_j(x)+x_ju_j\,\big)\boldsymbol{\Gamma}\big[\,dx\,\big].$$

Thus

$$\delta(u_1,\dots,u_n)=\sum_{j=1}^n\big(-\partial_{x_j}u_j(X_1,\dots,X_n)+X_ju_j(X_1,\dots,X_n)\,\big).$$

Observe that in the case $n=1$ the divergence operator coincides with the creation operator (3.1.2).

(b) Suppose that $\mathfrak{X}\subset L^2\big(\Omega,\mathcal{S},\mathbb{P}\big)$ is a separable Gaussian Hilbert space and $X\in\mathfrak{X}$. It is not hard to verify that $DX\in\operatorname{Dom}(\delta)$. We want to compute δDX.

For $F\in L^2_{\mathfrak{X}}(\Omega)$ we have

$$\mathbb{E}[F\delta DX]=\mathbb{E}\big[\,(Df,DX)_{\mathfrak{X}}\,\big],\;\;\forall\mathfrak{A}(\mathfrak{X}).$$

We can assume that $\|X\|_{L^2}=1$ and that $F=f(X_1,\dots,X_n)$, where $\{X_1,\dots,X_n\}$ is an orthonormal system, $X=X_1$ and f is an admissible function. We have

$$\mathbb{E}\big[\,(Df,DX)_{\mathfrak{X}}\,\big]=\int_{\mathbb{R}^n}f'_{x_1}(x)\boldsymbol{\Gamma}(dx)=\int_{\mathbb{R}^n}f(x)x_1\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dx\,\big]=\mathbb{E}[FX].$$

Hence

$$\delta(DX)=X,\;\;\forall X\in\mathfrak{X}.$$

(c) Suppose that $F\in\mathfrak{A}(\mathfrak{X})$ and $X\in\mathfrak{X}$. Then $D_XF=\big[\,DF,DX\,\big]_{\mathfrak{X}}\in\mathfrak{A}(X)$. Indeed, we can assume that $F=f(X_1,\dots,X_n)$, f admissible, $\{X_1,\dots,X_n\}$ orthonormal system, $X_1=X$. Then

$$\big[\,DF,DX\,\big]_{\mathfrak{X}}=f'_{x_1}(X_1,\dots,X_n)\in\mathfrak{A}(X).$$

Observe that

$$D(D_X F) = D\big[\,DF, DX\,\big]_{\mathfrak{X}} = \sum_{j=1}^{n} f''_{x_1 x_j}(X_1, \dots, X_n) DX_j = \frac{1}{2!} i_{X_1} D^2 F,$$

where for any $X \in \mathfrak{X}$ we denoted by i_X the contraction

$$i_X : \mathfrak{X}^{\otimes k} \to \mathfrak{X}^{\otimes (k-1)}, \;\; k \in \mathbb{N}$$

which is the $\otimes$-derivation uniquely determined by the condition

$$i_X Y = \big[\,X, Y\,\big]_{\mathfrak{X}}, \;\; \forall Y \in \mathfrak{X}. \qquad \square$$

The next result follows immediately from the definition of δ. We refer to [**124**, Prop. 2.5.4] for details.

Proposition 3.2.17. *Let* $F \in \mathbb{D}^{1,2}(\mathfrak{X})$ *and* $u \in \operatorname{Dom}(\delta)$ *such that*

$$\mathbb{E}\big[\,F^2 \|u\|^2_{\mathfrak{X}}\,\big] + \mathbb{E}\big[\,F^2 \delta(u)^2\,\big] + \mathbb{E}\big[\,\big[\,DF, u\,\big]^2_{\mathfrak{X}}\,\big] < \infty.$$

Then

$$\delta(Fu) = F\delta u - (DF, u)_{\mathfrak{X}}. \qquad \square$$

Example 3.2.18. Suppose that $F \in \mathfrak{A}(\mathfrak{X})$, $X \in \mathfrak{X}$ so that

$$u = FDX \in \mathfrak{A}(\mathfrak{X}, \mathfrak{X}).$$

Then $u \in \operatorname{Dom}(\delta)$ and we deduce from Proposition 3.2.17 that

$$\delta u = FX - (DF, DX)_{\mathfrak{X}} = FX - D_X F.$$

This shows that $\delta u \in \mathbb{D}^{1,2}$ and for any $Y \in \mathfrak{X}$ we have

$$D_Y(\delta u) = (D_Y F)X + F D_Y X - D_Y D_X F = (D_Y F)X + F\big[\,X, Y\,\big]_{\mathfrak{X}} - D_Y D_X F.$$

On the other hand

$$D_Y u = D_Y F \otimes DX, \;\; \delta(D_Y F u) = (D_Y F)X - D_X D_Y F.$$

Hence

$$\begin{aligned} D_Y(\delta u) - \delta(D_Y u) &= (D_Y F)X + F\big[\,X, Y\,\big]_{\mathfrak{X}} - D_Y D_X F - (D_Y F)X + D_X D_Y F \\ &= F(X, Y) + [D_X, D_Y]F = F(X, Y) = \big[\,u, Y\,\big]_{\mathfrak{X}}, \end{aligned}$$

where $[a, b]$ denotes the commutator of two elements a, b of an algebra. We have thus proved the *Heisenberg identity*

$$\forall u \in \operatorname{Dom}(\delta), \;\; Y \in \mathfrak{X} \;\; [D_Y, \delta]u = \big[\,u, Y\,\big]_{\mathfrak{X}}. \tag{3.2.10}$$

$\square$

The operator δ^p is closely related to the multiple Itô integrals. We have the following result.

Proposition 3.2.19. *Let* $X_1, \dots, X_p \in \mathfrak{X}$. *Then (compare with (3.1.48))*

$$\delta^p(DX_1 \otimes \cdots \otimes DX_p) = \delta^p\big(\,\mathbf{Sym}[DX_1 \otimes \cdots \otimes DX_p]\,\big) = \,:X_1 \cdots X_p:\,. \tag{3.2.11}$$

Proof. Fix an orthonormal basis $\{Y_n\}_{n\in\mathbb{N}}$ of $\mathfrak{X}$. Clearly it sufficers to prove the result in the special case when

$$DX_1 \odot \cdots \odot DX_p = (DY)^{\odot\alpha}, \;\; \alpha \in \mathbb{N}_0^{\mathbb{N}}, \;\; |\alpha| = p.$$

Suppose that $f = f(y_1, \dots, y_n)$ is an admissible function. Then

$$\mathbb{E}\big[\, f(Y_1, \dots, Y_n)\delta(DY^{\odot\alpha})\,\big] = \mathbb{E}\big[\,\big[\, D^p f(Y_1, \dots, Y_n), (DY)^{\odot\alpha}\,\big]_{\mathfrak{X}^{\odot p}}\,\big]$$

From Remark 3.2.4 we deduce

$$\mathbb{E}\big[\,\big[\, D^p f(Y_1, \dots, Y_n), (DY)^{\odot\alpha}\,\big]_{\mathfrak{X}^{\odot p}}\,\big] = \sqrt{p!}\sum_{|\beta|=p}\frac{1}{\beta!}\mathbb{E}\big[\,\partial_y^\beta f(Y_1, \dots, Y_n)\big[\, DY^{\odot\beta}, DY^{\odot\alpha}\,\big]_{\mathfrak{X}^{\odot p}}\,\big]$$

$$= \sqrt{p!}\mathbb{E}\big[\,\partial_y^\alpha f(Y_1, \dots, Y_n)\,\big] = \sqrt{p!}\int_{\mathbb{R}^n}\partial_y^\alpha f(y_1, \dots, y_n)\boldsymbol{\Gamma}_{\mathbb{1}}\big[\, dy\,\big]$$

$(\delta_{y_k} = -\partial_{y_k} + y_k\cdot,\; \delta_{y_k}^{\alpha_k}1 = H_{\alpha_k}(y_k))$

$$= \sqrt{p!}\int_{\mathbb{R}^n} f(y)H_\alpha(y)\boldsymbol{\Gamma}_{\mathbb{1}}\big[\, dy\,\big].$$

Hence

$$\delta^p((DY)^{\odot\alpha}) = \sqrt{|\alpha|!}H_\alpha(Y),$$

i.e.,

$$\delta^p\big(\,\mathbf{Sym}[(DY)^{\otimes\alpha}]\,\big) = \frac{1}{\sqrt{|\alpha|!}}\delta^p((DY)^{\odot\alpha}) = H_\alpha(Y) =: Y_1^{\alpha_1}\cdots Y_n^{\alpha_n} : .$$

This proves the second equality of (3.2.11). The first one is proved in a similar fashion. □

Remark 3.2.20. Using the equalities (3.1.34) and (3.1.41) we deduce that

$$\delta^p(u) = \sqrt{p!}\Theta_p(u), \;\; \forall p \in \mathbb{N}, \;\; \forall u \in \mathfrak{X}^{\odot p}.$$

If we are given a Hilbert space isomorphism

$$Z : L^2(M, \mathcal{M}, \mu) \to \mathfrak{X},$$

then the resulting map

$$L^2(M^p, \mu^{\otimes p}) \to \mathfrak{X}^{\otimes p} \xrightarrow{\delta^p} L^2_{\mathfrak{X}}\big(\,\Omega\,\big)$$

coincides with the multiple Wiener-Itô integral $\mathfrak{I}_n$; see (3.1.48). For this reason we set

$$\mathfrak{I}_p[F] := \delta^p F = \delta^p\big(\,\mathbf{Sym}[F]\,\big) = \sqrt{p!}\Theta_p\big(\,\mathbf{Sym}[F]\,\big), \;\; \forall F \in \mathfrak{X}^{\otimes p}. \tag{3.2.12}$$

Using the isometry relation (3.1.42) we deduce that

$$\mathbb{E}\big[\,\mathfrak{I}_p[F]^2\,\big] = \|\,\mathfrak{I}_p[F]\,\|^2 = p!\|F\,\|^2, \;\; \forall F \in \mathfrak{X}^{\odot p}. \tag{3.2.13}$$

□

Remark 3.2.21. For any Hilbert space H and any $k \in \mathbb{N}$ we have a Malliavin derivative

$$D_H^k : \mathfrak{A}(\mathfrak{X}, H) \to \mathfrak{A}(\mathfrak{X}, H \otimes \mathfrak{X}^{\otimes k})$$

with adjoint δ_H^k defined by the equality

$$\mathbb{E}\big[\, F\big[\, h, \delta(Gh' \otimes u)\,\big]_H\,\big] = \mathbb{E}\big[\,\big[\, D^k F \otimes h, Gh' \otimes u\,\big]_{H\otimes\mathfrak{X}^{\otimes k}}\,\big],$$

$\forall F, G \in \mathfrak{A}(\mathfrak{X})$, $h, h' \in H$, $u \in \mathfrak{X}^{\otimes k}$. For any $p \in \mathbb{N}$ we have

$$D^{p+1} = D_{\mathfrak{X}}^p \circ D, \;\; \delta^{p+1} = \delta_{\mathfrak{X}}^p \circ \delta.$$

Arguing as in the proof of Proposition 3.2.19 one can show

$$p\delta^{p-1}(u) = D\delta^p(u), \ \ \forall u \in \mathfrak{X} \otimes \mathfrak{X}^{\odot(p-1)}. \tag{3.2.14}$$

The above equality generalizes (3.1.4a). In fact, (3.2.14) follows from (3.1.4a). If as in the previous remark we set

$$\mathfrak{I}_{p-1}[u] = \delta^{p-1}(u), \ \ \forall u \in \mathfrak{X} \otimes \mathfrak{X}^{\odot(p-1)}.$$

We can rewrite (3.2.14) as

$$p\mathfrak{I}_{p-1}[u] = D\mathfrak{I}_p[u], \ \ \forall u \in \mathfrak{X} \otimes \mathfrak{X}^{\hat{\odot}(p-1)}. \tag{3.2.15}$$

□

3.2.3. The Ornstein-Uhlenbeck semigroup. Let $\mathfrak{X} \subset L^2(\Omega, \mathcal{S}, \mathbb{P})$ be a separable real Gaussian Hilbert space.

Definition 3.2.22. The *Ornstein-Uhlenbeck semigroup* is the semigroup of contractions $P_t : L^2_{\mathfrak{X}}(\Omega) \to L^2_{\mathfrak{X}}(\Omega)$, $t \geq 0$, defined by

$$T_t F = \sum_{n \geq 0} e^{-nt} \operatorname{Proj}_n F, \ \ \forall F \in L^2_{\mathfrak{X}}(\Omega), \ \ \forall t \geq 0,$$

where we recall that $\operatorname{Proj}_n : L^2_{\mathfrak{X}}(\Omega) \to \mathfrak{X}^{:n:}$ denotes the orthogonal projection onto the n-th chaos. □

The above definition shows that T_t is indeed a semigroup of selfadjoint L^2-contractions. It is a C_0-semigroup in the sense that

$$\lim_{t \searrow 0} T_t u = 0, \ \ \forall u \in L^2_{\mathfrak{X}}(\Omega).$$

We want to present an equivalent, coordinate dependent description of this semigroup.

Fix a complete *orthonormal* basis of $\mathfrak{X}$,

$$\underline{X} = (X_1, X_2, \dots, X_n, \dots).$$

Observe that the semigroup T_t is uniquely determined by its action on $\mathcal{P}(\mathfrak{X})$.

Proposition 3.2.23 (Mehler's formula)**.** *Let $P : \mathbb{R}^m \to \mathbb{R}$ be a polynomial in m real variables. Set*

$$\vec{X} := (X_1, \dots, X_m).$$

Then

$$T_t[P(\vec{X})](\omega) = \int_{\mathbb{R}^m} P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1 - e^{-2t}}y\Big)\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,y\,\big], \tag{3.2.16}$$

where $\boldsymbol{\Gamma}_{\mathbb{1}}$ denotes the canonical Gaussian measure on the Euclidean space $\mathbb{R}^m$.

Proof. It suffices to prove the result in the special case when $P(\vec{X}) = H_\alpha(\vec{X})$, $\alpha \in \mathbb{N}_0^m$. In this case, the left-hand side of (3.2.16) is equal to

$$T_t[P(\vec{X})](\omega) = e^{-|\alpha|t} H_\alpha(\vec{X}(\omega)) = \prod_{j=1}^m e^{-\alpha_j t} H_{\alpha_j}(X_j(t)).$$

The Fubini theorem shows that the right-hand side of (3.2.16) is equal in this case to

$$\prod_{j=1}^{m}\int_{\mathbb{R}} H_{\alpha_j}\Big(e^{-t}X_j(\omega)+\sqrt{1-e^{-2t}}\,y\Big)\boldsymbol{\Gamma}\big[\,dy\,\big].$$

Thus, to prove (3.2.16) it suffices to prove that

$$\int_{\mathbb{R}} H_n\Big(e^{-t}x++\sqrt{1-e^{-2t}}\,y\Big)\boldsymbol{\Gamma}\big[\,dy\,\big]=e^{-nt}H_n(x),\;\;\forall n\in\mathbb{N}_0,\;\;t\geq 0,\;\;\forall x\in\mathbb{R}. \tag{3.2.17}$$

We follow closely the presentation in the proof of [**94**, Prop. V.1.5.4]. We have the following useful identities.

Lemma 3.2.24. *Define the linear operator*

$$T_t:\mathbb{R}[x]\to\mathbb{R}[x],\;\;T_tP(x)=\int_{\mathbb{R}} P\Big(e^{-t}x++\sqrt{1-e^{-2t}}\,y\Big)\boldsymbol{\Gamma}\big[\,dy\,\big],$$

Then the following hold.

(i) *The operator T_t is symmetric with respect to the $L^2(\boldsymbol{\Gamma})$-inner product on $\mathbb{R}[x]$.*

(ii) $\partial_x T_t = e^{-t}T_t\partial_x$.

(iii) $T_t\delta_x = e^{-t}\delta_x T_t$.

Proof of Lemma 3.2.24. To prove (i) observe that

$$(T_tP,Q)=\int_{\mathbb{R}}\int_{\mathbb{R}} P\Big(e^{-t}x++\sqrt{1-e^{-2t}}\,y\Big)Q(x)\boldsymbol{\Gamma}\big[\,dy\,\big]\boldsymbol{\Gamma}\big[\,dx\,\big]. \tag{3.2.18}$$

Set $a=e^{-t}$, $b=\sqrt{1-e^{-2t}}$ so that $a^2+b^2=1$. We have

$$(T_tP,Q)=\int_{\mathbb{R}^2} P(ax+by)Q(x)\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dxdy\,\big].$$

Now consider the *orthogonal* change in variables

$$\begin{bmatrix} x\\ y\end{bmatrix}=\begin{bmatrix} b & a\\ -a & b\end{bmatrix}\cdot\begin{bmatrix} u\\ v\end{bmatrix}.$$

Since $\boldsymbol{\Gamma}_{\mathbb{1}}$ is invariant under orthogonal transformations we deduce

$$\int_{\mathbb{R}^2} P(ax+by)Q(x)\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dxdy\,\big]=\int_{\mathbb{R}^2} P(v)Q(av+bu)\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dudv\,\big]=(P,T_tQ).$$

This proves (i). The equality (ii) follows by differentiating the definition (3.2.18) of $T_t[P]$. The equality (iii) is obtained from (ii) by passing to adjoints, and using the symmetry of T_t proved in (i). □

Clearly, $T_t1=1$. From Lemma 3.2.24(iii) we deduce that

$$T_tH_n=T_t\delta_x^n1=e^{-nt}\delta_x^nT_t1=e^{-nt}H_n.$$

This concludes the proof of Proposition 3.2.23. □

The semigroup (P_t) is a C_0-semigroup of symmetric linear contractions on the Hilbert space $L^2_{\mathfrak{X}}(\Omega)$. According to Hille-Yosida's Theorem [130, Sec.1.3] P_t has the form $P_t = e^{tL}$, where L is a closed, densely defined, selfadjoint and nonpositive operator. Moreover, F is in the domain of L if and only if the limit

$$\lim_{t\searrow 0}\frac{1}{t}\Big(T_tF - F\Big)$$

exists in $L^2_{\mathfrak{X}}\big(\Omega\big)$. In this case, LF is the above limit.

Definition 3.2.25. The *Ornstein-Uhlenbeck operator* is the infinitesimal generator L of the Ornstein-Uhlenbeck semigroup. □

Proposition 3.2.26.

$$\mathrm{Dom}(L) = \Big\{\, F \in L^2_{\mathfrak{X}}\big(\Omega\big);\; \sum_{n\geq 0} n^2 \|\operatorname{Proj}_n\|^2_{L^2} < \infty \,\Big\} = \mathbb{D}^{2,2}(\mathfrak{X}).$$

$$\forall n \in \mathbb{N},\;\; \forall F \in \mathfrak{X}^{:n:}:\;\; LF = -nF = -\delta DF.$$

Proof. Let $F \in L^2_{\mathfrak{X}}\big(\Omega\big)$. We set $F_n = \operatorname{Proj}_n F$. Then

$$\frac{1}{t}\Big(T_tF - F\Big) = \sum_{n\geq 0}\frac{e^{-nt}-1}{t}F_n.$$

Now observe that

$$\Big|\frac{e^{-nt}-1}{t}\Big| \leq n,\;\; \forall t>0,\;\; N \in \mathbb{N}_0.$$

so that

$$\Big\|\frac{1}{t}\Big(T_tF - F\Big)\Big\|^2_{L^2} \leq \sum_{n\geq 0} n^2\|F_n\|^2_{L^2}$$

This proves that if

$$\sum_{n\geq 0} n^2\|F_n\|^2_{L^2} < \infty,$$

then

$$\lim_{t\searrow 0}\frac{1}{t}\Big(T_tF - F\Big)$$

exists in L^2 and it is equal to

$$\sum_{n\geq 0}\frac{d}{dt}\Big|_{t=0} e^{-nt}F_n = -\sum_{n\geq 0} nF_n.$$

Conversely, if the above limit exists in L^2, then

$$\operatorname{Proj}_n\left(\lim_{t\searrow 0}\frac{1}{t}\Big(T_tF - F\Big)\right) = \lim_{t\searrow 0}\operatorname{Proj}_n\left(\frac{1}{t}\Big(T_tF - F\Big)\right) = -nF_n.$$

Thus

$$\lim_{t\searrow 0}\frac{1}{t}\Big(T_tF - F\Big) = -\sum_{n\geq 0} nF_n \in L^2 \Rightarrow \sum_{n\geq 0} n^2\|F_n\|^2_{L^2} < \infty.$$

The equality $LF = -nF$, $f \in \mathfrak{X}^{:n:}$ follows from the above discussion. To prove the equality $\delta DF = n$, $F \in \mathfrak{X}^{:n:}$ it suffices to consider only the special case when $F = H_\alpha(X_1, \dots, X_k)$

where (X_j) is an orthonormal system and α is a multi-index such that $|\alpha| = n$. In this case the equality follows from (3.1.4b). □

Example 3.2.27. (a) Suppose that $\dim \mathfrak{X} = n$. By fixing an orthonormal basis $X_1, \dots, X_n$ of $\mathfrak{X}$ we can identify $L^2_{\mathfrak{X}}\big(\Omega\big)$ with $L^2(\mathbb{R}^n, \boldsymbol{\Gamma}_{\mathbb{1}})$. Then

$$Lf = \sum_{j=1}^{n} \partial^2_{x_j} f - \sum_{j=1}^{n} x_j \partial_{x_j} f = (-\Delta - x\nabla)f,$$

for any function$f \in C^2(\mathbb{R}^n)$ with bounded 2nd order derivatives. Above, Δ is the Euclidean *geometers'* Laplacian. In particular, Δ is nonnegative. □

Definition 3.2.28. We define L^{-1} to be the bounded operator $L^{-1} : L^2_{\mathfrak{X}}\big(\Omega\big) \to L^2_{\mathfrak{X}}\big(\Omega\big)$ given by

$$L^{-1}F = -\sum_{n\geq 1} \frac{1}{n} \operatorname{Proj}_n F. \qquad \square$$

Note that L^{-1} is a pseudo-inverse of L. More precisely, if $F \in \mathbb{D}^{2,2}(\mathfrak{X})$ is such that $\mathbb{E}\big[\, F \,\big] = 0$, i.e., $\operatorname{Proj}_0 F = 0$, then

$$L^{-1}LF = LL^{-1}F = F.$$

Proposition 3.2.29. *Let $F \in \mathbb{D}^{1,2}(\mathfrak{X})$. Then for any $X \in \mathfrak{X}$, $\|X\|_{L^2} = 1$, we have*

$$D_X L^{-1}F = -\int_0^\infty e^{-t} T_t D_X F dt = (L - \mathbb{1})^{-1} D_X F. \tag{3.2.19}$$

Proof. It suffices to prove the result in the special case

$$F = H_\alpha(X_1, \dots, X_m),$$

where $\{X_1, \dots, X_m\} \subset \mathfrak{X}$ is an orthonormal system, $X = X_1$, $|\alpha| = n > 0$. Note that

$$D_X F = \alpha_1 H_\beta(X_1, \dots, X_m), \quad \beta = (\alpha_1 - 1, \alpha_2, \dots, \alpha_m).$$

Using the identity

$$\frac{1}{n} = \int_0^\infty e^{-nt} dt$$

we deduce

$$L^{-1}F = -\frac{1}{n} F = -\int_0^\infty T_t F dt \Rightarrow D_X L^{-1} F = -\int_0^\infty D_X T_t F dt = -\int_0^\infty e^{-t} T_t D_X F dt.$$

On the other hand

$$(L - \mathbb{1})^{-1} D_X F = (L - \mathbb{1})^{-1}[\alpha_1 H_\beta] = -\frac{1}{|\beta| + 1} \alpha_1 H_\beta = -\frac{1}{n} D_X F = D_X L^{-1} F.$$

□

Proposition 3.2.30 (Key integration by parts formula)**.** *Suppose that $F, G \in \mathbb{D}^{1,2}(\mathfrak{X})$ are non-constant and $g : \mathbb{R} \to \mathbb{R}$ is a C^1-function with bounded derivative. Then*

$$\mathbb{E}\big[\, Fg(G) \,\big] = \mathbb{E}[F] \cdot \mathbb{E}\big[\, g(G) \,\big] + \mathbb{E}\big[\, g'(G) \cdot \big[\, DG, -DL^{-1}F \,\big]_{\mathfrak{X}} \,\big]. \tag{3.2.20}$$

Proof. Let $F_\perp = F - \mathbb{E}[F]$. Then

$$\mathbb{E}\big[\, Fg(G)\,\big] = \mathbb{E}[F]\cdot\mathbb{E}\big[\, g(G)\,\big] + \mathbb{E}\big[\, F_\perp g(G)\,\big].$$

Since $F_\perp = LL^{-1}F$ we deduce

$$\mathbb{E}\big[\, F_\perp g(G)\,\big] = \mathbb{E}\big[\, LL^{-1}Fg(G)\,\big] = -\mathbb{E}\big[\, \delta DL^{-1}Fg(G)\,\big]$$

$$\stackrel{(3.2.9)}{=} -\mathbb{E}\big[\,\big[\, DL^{-1}F, Dg(G)\,\big]_{\mathfrak{X}}\,\big] \stackrel{(3.2.8)}{=} -\mathbb{E}\big[\,\big[\, DL^{-1}F, g'(G)DG\,\big]_{\mathfrak{X}}\,\big]$$

$$= \mathbb{E}\big[\, g'(G)\big[\, DG, -DL^{-1}F\,\big]_{\mathfrak{X}}\,\big].$$

□

3.2.4. The hyper-contractivity of the Ornstein-Uhlenbeck semigroup. We know that (T_t) defines a C_0-semigroup of contractions $L^2_{\mathfrak{X}}\big(\,\Omega\,\big) \to L^2_{\mathfrak{X}}\big(\,\Omega\,\big)$.

Proposition 3.2.31. *For any $t > 0$ and any $p \in (1,\infty)$ the operator T_t defines a contraction $L^p_{\mathfrak{X}}\big(\,\Omega\,\big) \to L^p_{\mathfrak{X}}\big(\,\Omega\,\big)$.*

Proof. We limit ourself to proving that

$$\|T_tP\|_{L^p} \le \|P\|_{L^p},\ \ \forall P \in \mathcal{P}(\mathfrak{X}).$$

To see this assume $P = P(\vec{X})$, where $\vec{X} = (X_m)_{m\ge1}$ is an orthonormal system in $\mathfrak{X}$. Using Mehler's formula (3.2.16) we deduce

$$T_t[P(\vec{X})](\omega) = \int_{\mathbb{R}^m} P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1-e^{-2t}}y\Big)\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dy\,\big].$$

Since the function $f(x) = x^p$, $x > 0$, is convex for $p > 1$ we deduce from Jensen's inequality that

$$\big|\, T_t[P(\vec{X})](\omega)\,\big|^p \le \int_{\mathbb{R}^m}\Big|\, P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1-e^{-2t}}\,y\Big)\,\Big|^p\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dy\,\big].$$

Invoking Jensen's inequality once again we conclude that

$$\mathbb{E}\big[\,|T_tP|^p\,\big] \le \int_{\mathbb{R}^m}\mathbb{E}\Big[\Big|\, P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1-e^{-2t}}\,y\Big)\,\Big|^p\Big]\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dy\,\big]$$

$$= \int_{\mathbb{R}^m\times\mathbb{R}^m}\Big|\, P(e^{-t}x + \sqrt{1-e^{-2t}}y\,)\Big|^p\boldsymbol{\Gamma}_{\mathbb{1}}(dx)\boldsymbol{\Gamma}_{\mathbb{1}}(dy) = \int_{\mathbb{R}^m}\big|\, P(\,x\,)\,\big|^p\boldsymbol{\Gamma}_{\mathbb{1}}\big[\,dx\,\big],$$

where at the last step we used the fact that if X, Y are independent standard normal random variables and $a^2 + b^2 = 1$, then $aX + bY$ is also a standard normal random variable. □

The semigroup T_t satisfies a hypercontractivity property, namely, for any $p_0 \in (1,\infty)$ there exists a strictly increasing, unbounded function $p : [0,\infty) \to (0,\infty)$ such that $p_0 = p(0)$ and, $\forall t \ge 0$, the operator T_t induces a bounded linear map $T_t : L^{p_0} \to L^{p(t)}$. We will spend the remainder of this subsection proving this fact. We denote by $\boldsymbol{\Gamma}_{\mathbb{1}}$ the canonical Gaussian measure on a finite dimensional Euclidean space.

Theorem 3.2.32 (The log-Sobolev inequality)**.** *For any $n \in \mathbb{N}$, and any $f \in W^{1,2}(\mathbb{R}^n, \mathbf{\Gamma}_{\mathbb{1}})$ we have*

$$\begin{aligned}\int_{\mathbb{R}^n} f^2(x) \log f^2(x) \mathbf{\Gamma}_{\mathbb{1}}\big[\,dx\,\big] \leq 2 \int_{\mathbb{R}^n} |\nabla f(x)|^2 \,\mathbf{\Gamma}_{\mathbb{1}}\big[\,dx\,\big] \\ + \int_{\mathbb{R}^n} f^2(x) \mathbf{\Gamma}_{\mathbb{1}}\big[\,dx\,\big] \log\left(\int_{\mathbb{R}^n} f^2(x)\, \mathbf{\Gamma}_{\mathbb{1}}\big[\,dx\,\big]\right),\end{aligned} \tag{3.2.21}$$

where $0 \cdot \log 0 := 0$.

Proof. We follow the presentation in [**21**, §1.6]. Assume first that $f \in C_b^\infty(\mathbb{R}^n)$, i.e., f and all its derivatives are bounded. We distinguish three cases.

A. $\exists c > 0$ such that $f(x) > c$, $\forall x \in \mathbb{R}^n$. Set $\varphi = f^2$ so that

$$\nabla f = \frac{1}{2\sqrt{\varphi}} \nabla \varphi$$

and (3.2.21) is equivalent to

$$\int_{\mathbb{R}^n} \varphi \log \varphi \, d\mathbf{\Gamma}_{\mathbb{1}} - \int_{\mathbb{R}^n} \varphi d\mathbf{\Gamma}_{\mathbb{1}} \log\left(\int_{\mathbb{R}^n} \varphi \, d\mathbf{\Gamma}_{\mathbb{1}} \right) \leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{1}{\varphi} |\nabla \varphi|^2 \, d\mathbf{\Gamma}. \tag{3.2.22}$$

Consider the Ornstein-Uhlenbeck semigroup

$$T_t : L^2(\mathbb{R}^n, \mathbf{\Gamma}_{\mathbb{1}}) \to L^2(\mathbb{R}^n, \mathbf{\Gamma}_{\mathbb{1}}).$$

Using the equality (3.2.16) we deduce that

$$T_t[\varphi](x) \geq c, \;\; \forall x \in \mathbb{R}^n, \;\; t \geq 0.$$

Since

$$\lim_{t\to\infty} T_t[\varphi] \log T_t[\varphi] = \int_{\mathbb{R}^n} \varphi \, d\mathbf{\Gamma}_{\mathbb{1}} \log\left(\int_{\mathbb{R}^n} \varphi \, d\mathbf{\Gamma}_{\mathbb{1}} \right)$$

we see that the left-hand side of (3.2.22) is equal to

$$-\int_0^\infty \frac{d}{dt} \int_{\mathbb{R}^n} T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}|one.$$

Taking into account the fact that

$$\frac{d}{dt} T_t[g] = L T_t[g], \;\; \forall g \in C_b^\infty(\mathbb{R}^n)$$

we deduce

$$\begin{aligned} &-\int_0^\infty \frac{d}{dt} \int_{\mathbb{R}^n} T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}_{\mathbb{1}} \\ &= -\int_0^\infty \int_{\mathbb{R}^n} L T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}_{\mathbb{1}} - \int_0^\infty \int_{\mathbb{R}^n} T_t[\varphi] \frac{1}{T_t[\varphi]} \frac{d}{dt} T_t[\varphi] \, d\mathbf{\Gamma}_{\mathbb{1}} \\ &= -\int_0^\infty \int_{\mathbb{R}^n} L T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}_{\mathbb{1}} - \int_0^\infty \int_{\mathbb{R}^n} L T_t[\varphi] \, d\mathbf{\Gamma}_{\mathbb{1}}. \end{aligned}$$

Since L is symmetric and $L1 = 0$ we deduce

$$\int_{\mathbb{R}^n} L T_t[\varphi] \, d\mathbf{\Gamma}_{\mathbb{1}} = 0.$$

Hence

$$\int_{\mathbb{R}^n} \varphi \log \varphi \, d\mathbf{\Gamma}_{\mathbb{1}} - \int_{\mathbb{R}^n} \varphi d\mathbf{\Gamma} \log\left(\int_{\mathbb{R}^n} \varphi \, d\mathbf{\Gamma}_{\mathbb{1}} \right) = -\int_0^\infty \int_{\mathbb{R}^n} L T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}_{\mathbb{1}}$$

$$= \int_0^\infty \int_{\mathbb{R}^n} \delta D T_t[\varphi] \log T_t[\varphi]\, d\mathbf{\Gamma}_{\mathbb{1}} = \int_0^\infty \int_{\mathbb{R}^n} \big(\nabla T_t[\varphi], \nabla \log T_t[\varphi]\,\big)\, d\mathbf{\Gamma}_{\mathbb{1}}$$

$$= \int_0^\infty \underbrace{\int_{\mathbb{R}^n} \frac{1}{T_t[\varphi]} \big|\,\nabla T_t[\varphi]\,\big|^2 d\mathbf{\Gamma}_{\mathbb{1}}}_{F(t)}\,.$$

Using Lemma 3.2.24 (ii) we deduce

$$\partial_{x_i} T_t[\varphi] = e^{-t} T_t[\partial_{x_i}\varphi], \ \ \forall i = 1, \dots, n,$$

so that

$$F(t) = e^{-2t} \int_{\mathbb{R}^n} \frac{1}{T_t[\varphi]} \sum_{i=1}^n \big(\,T_t[\partial_{x_i}\varphi]\big)^2 d\mathbf{\Gamma}_{\mathbb{1}}.$$

The equality (3.2.16) implies that for any $g, h \in C_b^\infty(\mathbb{R}^n)$ we have

$$T_t[g] \le T_t[\,|g|\,] \le \|g\|_{L^\infty}, \ \ \big(\,T_t[gh]\,\big)^2 \le T_t[g^2] T_t[h]^2.$$

Hence

$$\frac{1}{T_t[\varphi]} \big(T_t[\partial_{x_i}\varphi]\big)^2 = \frac{1}{T_t[\varphi]} \left(T_t \left[\sqrt{\varphi} \cdot \frac{\partial_{x_i}\varphi}{\sqrt{\varphi}} \right] \right)^2 \le T_t \left[\frac{(\partial_{x_i}\varphi)^2}{\varphi} \right] \le \frac{(\partial_{x_i}\varphi)^2}{\varphi}.$$

Thus

$$F(t) \le e^{-2t} \int_{\mathbb{R}^n} \frac{|\nabla\varphi|^2}{\varphi} d\mathbf{\Gamma}_{\mathbb{1}}.$$

The inequality (3.2.22) follows by integrating the above inequality.

B. $f \in W^{1,2}(\mathbb{R}^n, \mathbf{\Gamma})$, $f \ge 0$ a.s. . This case follows from case **A** by choosing a family of functions $f_\varepsilon \in C_b^\infty(\mathbb{R}^n)$, $f_\varepsilon \ge \varepsilon$, $f_\varepsilon \to f$ in $W^{1,2}$ and then letting $\varepsilon \searrow 0$.

The general case, $f \in W^{1,2}(\mathbb{R}^n, \mathbf{\Gamma}_{\mathbb{1}})$, follows from case **B** applied to $|f|$. □

Remark 3.2.33. If $(\Omega, \mathcal{S}, \mu)$ is a probability space and $f : \Omega \to [0, \infty)$ is measurable function, then its entropy with respect to μ is

$$\operatorname{Ent}_\mu[f] = \begin{cases} \mathbb{E}_\mu[f \log f] - \mathbb{E}_\mu[f] \log \mathbb{E}_\mu[f], & \mathbb{E}_\mu[\log(1+f)] < \infty, \\ \\ +\infty, & \mathbb{E}_\mu[\log(1+f)] = \infty. \end{cases}$$

where $0 \log 0 := 0$. Observe that $\operatorname{Ent}_\mu(f)$ is nonnegative and positively homogeneous of degree 1. The log-Sobolev inequality (3.2.21) can be rewritten as

$$\operatorname{Ent}_{\mathbf{\Gamma}_1}[f^2] \le 2 \int_{\mathbb{R}^n} |\nabla f(x)|^2 \mathbf{\Gamma}_{\mathbb{1}}\big[\,dx\,\big].$$

As explained in [**87**, Sec.5.1], the log-Sobolev inequality leads to rather sharp concentration of measure inequalities. □

Theorem 3.2.34 (Hypercontractivity)**.** *Let $p \in (1, \infty)$. Define*

$$q(t) := 1 + e^{2t}(p-1), \ \ \forall t \ge 0.$$

Then

$$\|T_t f\|_{L^{q(t)}} \le \|f\|_{L^p}, \ \ \forall f \in L^p(\mathbb{R}^n, \mathbf{\Gamma}_{\mathbb{1}}), \ \ t \ge 0. \tag{3.2.23}$$

Note that $q(t) > p$, $\forall t > 0$.

Proof. We follow closely the arguments in [**21**, Thm. 1.6.2]. It suffices to prove the inequality for smooth functions $f \in C_b^\infty(\mathbb{R}^n)$ such that

$$c := \inf_{x \in \mathbb{R}^n} f(x) > 0.$$

Under this assumption the function $[0,\infty) \ni t \mapsto G(t) = \|f\|_{L^{q(t)}}$ is differentiable. The inequality (3.2.23) reads $G(t) \leq G(0)$ so it it suffices to prove that $G'(t) \leq 0$, $\forall t \geq 0$.

Applying the log-Sobolev inequality to the function $f^{r/2}$, $r > 0$, we deduce

$$\int_{\mathbb{R}^n} f^r \log f d\boldsymbol{\Gamma}_{\mathbb{1}} - \frac{1}{r}\int_{\mathbb{R}^n} f^r d\boldsymbol{\Gamma}_{\mathbb{1}}\left(\log \int_{\mathbb{R}^n} f^r d\boldsymbol{\Gamma}_{\mathbb{1}}\right) \leq \frac{r}{2}\int_{\mathbb{R}^n}\left(f^{r-2}\nabla f, \nabla f\right) d\boldsymbol{\Gamma}_{\mathbb{1}}$$

$$= \frac{r}{2(r-1)}\int_{\mathbb{R}^n}(\nabla f^{r-1}, \nabla f) d\boldsymbol{\Gamma}_{\mathbb{1}} = -\frac{r}{2(r-1)}\int_{\mathbb{R}^n} f^{r-1} L f d\boldsymbol{\Gamma}_{\mathbb{1}}.$$

Hence, $\forall r > 0$,

$$\int_{\mathbb{R}^n} f^r \log f d\boldsymbol{\Gamma}_{\mathbb{1}} - \frac{1}{r}\int_{\mathbb{R}^n} f^r d\boldsymbol{\Gamma}_{\mathbb{1}}\left(\log \int_{\mathbb{R}^n} f^r d\boldsymbol{\Gamma}_{\mathbb{1}}\right) \leq -\frac{r}{2(r-1)}\int_{\mathbb{R}^n} f^{r-1} L f d\boldsymbol{\Gamma}_{\mathbb{1}}. \tag{3.2.24}$$

We set

$$F(t) := \int_{\mathbb{R}^n} T_t[f]\, d\boldsymbol{\Gamma}_{\mathbb{1}}.$$

Then $G(t) = F(t)^{1/q(t)}$ and we have

$$G'(t) = G(t)\left(-\frac{q'(t)}{q(t)^2}\log F(t) + \frac{F'(t)}{q(t)F(t)}\right).$$

Since $q'(t) = 2q(t) - 2 > 0$ it suffices to show that

$$-\frac{1}{q(t)}F(t)\log F(t) + \frac{F'(t)}{q(t)} \leq 0. \tag{3.2.25}$$

Observing that

$$F'(t) = \int_{\mathbb{R}^n}(T_t[f])^{q(t)}\left(q'(t)\log T_t[f] + q(t)\frac{LT_t[f]}{T_t[f]}\right) d\boldsymbol{\Gamma}$$

we conclude that (3.2.24) is equivalent to

$$-\frac{F(t)\log F(t)}{q(t)} + \int_{\mathbb{R}^n}\big(T_t[f]\big)^{q(t)}\log T_t[f] d\boldsymbol{\Gamma}_{\mathbb{1}} + \frac{q(t)}{q'(t)}\int_{\mathbb{R}^n}\big(T_t[f]\big)^{q(t)-1} LT_t[f] d\boldsymbol{\Gamma}_{\mathbb{1}} \leq 0.$$

This is precisely the inequality (3.2.24) with $r = q(t)$. □

Corollary 3.2.35. *Let $\mathfrak{X} \subset L^2\big(\Omega, \mathcal{S}, \mathbb{P}\big)$ be a separable Gaussian Hilbert space. Let $p > 1$, Then for any $t > 0$, $F \in L^2_{\mathfrak{X}}\big(\Omega\big]$ we have*

$$\|T_t F\|_{L^{q(t)}} \leq \|F\|_{L^p}, \;\; \forall F \in L^p\mathfrak{X}\big(\Omega\big), \;\; q(t) = 1 + e^{2t}p. \tag{3.2.26}$$

Proof. Follows from Theorem 3.2.34 and the density of $\mathcal{P}(\mathfrak{X})$ is dense in $L^p_{\mathfrak{X}}(\Omega)$. □

Corollary 3.2.36. *Let $n \in \mathbb{N}$ and $F \in \mathfrak{X}^{:n:} \subset L^2_{\mathfrak{X}}\big(\Omega\big)$. Then $F \in \mathfrak{X}^{:n:} \subset L^q_{\mathfrak{X}}\big(\Omega\big)$, $\forall q \in [1,\infty)$.*

Proof. The claim is obviously true for $q \in [1,2]$. Assume that $q > 2$ Note that $T_t F = e^{-nt}F$. On the other hand $T_t F \in L^{1+e^{2t}}_{\mathfrak{X}}\big(\Omega\big)$ for any $t < 0$. Hence, if $1 + e^{2t} > q$, then $e^{nt}F \in L^q$. □

We conclude by mentioning, without proof, the Kree-Meyer inequality.

Theorem 3.2.37 (Kree-Meyer)**.** *For any* $p \in (1,\infty)$*, and any* $k,\ell \in \mathbb{N}_0$*, there exist positive constants* $c_p(k,\ell) < C_p(k,\ell)$ *such that*

$$c_p\|F\|_{\mathbb{D}^{k+\ell,p}} \leq \|(\mathbb{1}-L)^{\frac{\ell}{2}}F\|_{\mathbb{D}^{k,p}} \leq C_p\|F\|_{\mathbb{D}^{k+\ell,p}}, \quad \forall F \in \mathfrak{A}(\mathfrak{X}). \tag{3.2.27}$$

□

For a proof we refer to [**21**, Sec. 5.6], [**94**, Chap 2] or [**126**, Sec. 1.5].

3.3. The Stein method

3.3.1. Metrics on spaces of probability measures. Let us recall several concepts of of pseudo-distances on the spaces of Borel probability measures on $\mathbb{R}^d$.

Definition 3.3.1. Let $\mathcal{H}$ be a set of Borel measurable functions $\mathbb{R}^d \to \mathbb{R}$. We denote by $\operatorname{Prob}\big(\mathbb{R}^d\big)$ the space of Borel probability measures on $\mathbb{R}^d$.

(i) We set

$$\operatorname{Prob}\big(\mathbb{R}^d,\mathcal{H}\big) := \big\{\, \mu \in \operatorname{Prob}\big(\mathbb{R}^d\big);\;\; \mathcal{H} \subset L^1(\mathbb{R}^d,\mu)\,\big\}.$$

(ii) We say that $\mathcal{H}$ is called *separating* if for any $\mu,\nu \in \operatorname{Prob}\big(\mathbb{R}^d\big)$

$$\mu = \nu \Longleftrightarrow \mathbb{E}_\mu\big[\,h\,\big] = \mathbb{E}_\nu\big[\,h\,\big],\;\; \forall h \in \mathcal{H} \cap L^1\big(\mathbb{R}^d,\,\mu\big) \cap L^1\big(\mathbb{R}^d,\,\nu\big).$$

(iii) If $\mathcal{H}$ is separating and $\mu,\nu \in \operatorname{Prob}\big(\mathbb{R}^d,\mathcal{H}\big)$, we set

$$\operatorname{dist}_{\mathcal{H}}(\mu,\nu) := \sup_{h\in\mathcal{H}}\big|\,\mathbb{E}_\mu\big[\,h\,\big] - \mathbb{E}_\nu\big[\,h\,\big]\,\big|.$$

(iv) If $(\Omega,\mathcal{S},\mathbb{P})$ is a probability space and $F,G:\Omega\to\mathbb{R}^d$ are random variables whose probability distributions belong to $\operatorname{Prob}\big(\mathbb{R}^d,\mathcal{H}\big)$, then we set

$$\operatorname{dist}_{\mathcal{H}}(F,G) := \operatorname{dist}_{\mathcal{H}}\big(\mathbb{P}_F,\mathbb{P}_G\big) = \sup_{h\in\mathcal{H}}\big|\,\mathbb{E}\big[\,h(F)\,\big] - \mathbb{E}\big[\,h(G)\,\big]\,\big|.$$

□

It is easy to check that if $\mathcal{H}$ is separating, then $\operatorname{dist}_{\mathcal{H}}$ is indeed a metric on $\operatorname{Prob}\big(\mathbb{R}^d,\mathcal{H}\big)$.

Example 3.3.2. (a) If $\mathcal{H}$ is the class of functions

$$\boldsymbol{I}_{(-\infty,c_1]\times\cdots\times(-\infty,c_d]},\quad c_1,\dots,c_d \in \mathbb{R},$$

then the resulting metric $\operatorname{dist}_{\mathcal{H}}$ on $\operatorname{Prob}\big(\mathbb{R}^d\big)$ is called the *Kolmogorov distance* and it is denoted by $\operatorname{dist}_{Kol}$.

(b) If $\mathcal{H}$ is the class of bounded Borel measurable functions $h:\mathbb{R}^d\to[0,1]$, then $\mathcal{H}$ is separating then the resulting metric on $\operatorname{Prob}\big(\mathbb{R}^d\big)$ is called the *total variation* metric and it is denoted by dist_{TV}.

(c) If $\mathcal{H}$ is the class of Lipschitz continuous functions $\mathbb{R}^d\to\mathbb{R}$ satisfying $\operatorname{Lip}(h)\leq 1$, where $\operatorname{Lip}(h)$ is the (best) Lipschitz constant of h, then $\mathcal{H}$ is separating, the resulting metric is called the *Wasserstein* metric and it is denoted by dist_W.

(d) If $\mathcal{H}$ denotes the class of Lipschitz continuous functions $h : \mathbb{R}^d \to \mathbb{R}$ such that

$$\|h\|_{L^\infty} + \mathrm{Lip}(h) \leq 1,$$

then $\mathcal{H}$ is separating, the resulting distance is called the *Fortet-Mourier* metric and it is denoted by dist_{FM}.

(e) If $\mathcal{H} \subset C_b^2(\mathbb{R}^d)$ denotes the class of C^2-functions $f : \mathbb{R}^d \to \mathbb{R}$ satisfying

$$\|f\|_{C^2} \leq 1,$$

then $\mathcal{H}$ is separating. We denote by dist_{C^2} the resulting metric on $\mathrm{Prob}\big(\mathbb{R}^d\big)$. □

Remark 3.3.3. (a) Clearly

$$\mathrm{dist}_{Kol} \leq \mathrm{dist}_{TV}, \;\; \mathrm{dist}_{FM} \leq \mathrm{dist}_{W}, \;\; \mathrm{dist}_{C^2} \leq \mathrm{dist}_{W}$$

Thus

$$\lim_{n\to\infty} \mathrm{dist}_{TV}(F_n, F) = 0 \Rightarrow \lim_{n\to\infty} \mathrm{dist}_{Kol}(F_n, F) = 0.$$

Moreover, if

$$\lim_{n\to\infty} \mathrm{dist}_{Kol}(F_n, F) = 0,$$

then $F_n \to F$ in law.

(b) Also, one can prove (see [**49**, Thm.11.3.3] that $F_n \to F$ in distribution of and only if $F_n \to F$ in the Fortet-Mornier metric. It is not hard to see that dist_{C^2} induces on $\mathrm{Prob}\big(\mathbb{R}^d\big)$ the same topology as dist_{FM}, the topology of convergence in law. Moreover, (see [**34**, Thm.3.3]), if $N \sim \mathcal{N}(0,1)$, then

$$\mathrm{dist}_{Kol}(F, N) \leq 2\sqrt{\mathrm{dist}_W(F, N)}.$$ □

The Stein method provides a way of estimating the distance between a random variable and a normal random variable. I will present the bare-bones minimum referring to [**33, 34, 137**] for more details and many more applications. For more recent developments I refer to [**14, 90**]. I am following the presentation in [**124**, Chap.3,4]. It all starts with the following simple observation.

3.3.2. The one-dimensional Stein method. Suppose that $N \sim \mathcal{N}(0,1)$ and $g \in \mathbb{D}^{1,2}(\mathbb{R})$, i.e., $g(N), g'(N) \in L^2$. Then

$$\int_{\mathbb{R}} \big(-g'(x) + xg(x) \big) \boldsymbol{\Gamma}_1(dx) = \int_{\mathbb{R}} \delta_x g(x) \cdot 1 \, \boldsymbol{\Gamma}_1(dx) = \int_{\mathbb{R}} g(x) \cdot (\partial_x 1) \boldsymbol{\Gamma}_1(dx) = 0,$$

so that

$$\mathbb{E}\big[\, Ng(N)\,\big] = \mathbb{E}\big[\, g'(N)\,\big], \;\; \forall g \in \mathbb{D}^{1,2}(\mathbb{R}). \tag{3.3.1}$$

It turns out that the converse is also true.

Lemma 3.3.4 (Stein's Lemma)**.** *A random variable X is a standard normal random variable if and only if for all $g \in C^1(\mathbb{R})$ such that $g' \in L^1\big(\mathbb{R}, \boldsymbol{\Gamma}_1\big)$ and*

$$\mathbb{E}\big[\, Xg(X)\,\big] = \mathbb{E}\big[\, g'(X)\,\big]. \tag{3.3.2}$$

Proof. The necessity follows from Proposition 1.1.4. To prove the sufficiency note that ussing (3.3.2) with $g(x) = x^k$, $k = 0, 1, \dots$, we deduce

$$\mathbb{E}\big[\,X^{k+1}\,\big] = k\mathbb{E}\big[\,X^{k-1}\,\big],\ \ \forall k = 0, 1, 2, \dots .$$

This proves $\mathbb{E}\big[\,X\,\big] = 0$, $\mathbb{E}\big[\,X^2\,\big] = 1$ and that

$$\mathbb{E}\big[\,X^n\,\big] = \begin{cases} (n-1)!!, & n \equiv 0 \bmod 2,\\ 0, & n \equiv 1 \bmod 2 \end{cases} = \int_{\mathbb{R}} x^n \boldsymbol{\Gamma}_1(dx),\ \ \forall n = 1, 2 \dots,$$

The conclusion follows from the fact that the normal distribution is uniquely determined by its moments. □

Stein's lemma suggests that for a random variable X the quantity $\mathbb{E}[Xf(X) - f'(X)\,\big]$ should give an indication of how far away is the distribution of X from the normal distribution.

Definition 3.3.5. Let $N \sim \mathcal{N}(0,1)$ and $h \in L^2(\mathbb{R}, \boldsymbol{\Gamma}_1)$. The *Stein's equation* associated to h is the o.d.e.

$$\underbrace{g'(x) - xg(x)}_{=-\delta_x g(x)} = h(x) - \int_{\mathbb{R}} h(x)\boldsymbol{\Gamma}_1(dx) = h(x) - \mathbb{E}[h(N)]. \tag{3.3.3}$$

We set $h_\perp(x) := h(x) - \mathbb{E}[h(N)]$ so that

$$\mathbb{E}[\,h_\perp(N)\,] = 0.$$

□

Observe that Stein's equation can be rewritten as

$$e^{\frac{x^2}{2}}\partial_x\big(\,e^{-\frac{x^2}{2}}g(x)\,\big) = h_\perp(x), \tag{3.3.4}$$

If g_1, g_2 are two solutions of the linear equation (3.3.4), then

$$e^{x^2/2}\big(\,g_1(x) - g_2(x)\,\big) = \text{constant}.$$

This implies immediately the following result.

Proposition 3.3.6. *The general solution of (3.3.3) has the form*

$$g(x) = g_{h,c}(x) = ce^{\frac{x^2}{2}} + e^{\frac{x^2}{2}}\int_{-\infty}^{x} h_\perp(y)e^{-\frac{y^2}{2}}dy,\ \ x \in \mathbb{R}, \tag{3.3.5}$$

where $c \in \mathbb{R}$ is an arbitrary real constant. Moreover the solution

$$g_h(x) := g_{h,c=0} = e^{\frac{x^2}{2}}\int_{-\infty}^{x} h_\perp(y)e^{-\frac{y^2}{2}}dy \tag{3.3.6}$$

is the unique solution $g(x)$ of (3.3.3) such that

$$\lim_{x\to\pm\infty} e^{-\frac{x^2}{2}}g(x) = 0. \tag{3.3.7}$$

□

If now F is a random variable, then integrating the equality

$$g_h'(x) - xg_h(x) = h(x) - \mathbb{E}[\,h(N)\,]$$

with respect to the probability distribution of F we deduce

$$\mathbb{E}[\,h(F)\,] - \mathbb{E}[\,h(N)\,] = \mathbb{E}\big[\,g_h'(F) - Fg_h(F)\,\big]. \tag{3.3.8}$$

Thus, if $\mathcal{H}$ is a separating collection of Borel measurable functions $h : \mathbb{R} \to \mathbb{R}$ we deduce

$$\operatorname{dist}_{\mathcal{H}}(F, N) = \sup_{h \in \mathcal{H}} \Big| \mathbb{E}\big[g_h'(F) - F g_h(F) \big] \Big|. \tag{3.3.9}$$

We want to use the above equality to produce estimates on the Wasserstein distance between two Borel probability measures on $\mathbb{R}$.

Proposition 3.3.7. *Let $h : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function. Set $K := \operatorname{Lip}(h)$. Then the function g_h given by (3.3.6) admits the representation*

$$g_h(x) = -\int_0^\infty \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \mathbb{E}\big[h\big(e^{-t}x + \sqrt{1-e^{-2t}} N \big) N \big] dt. \tag{3.3.10}$$

Moreover, g_h is a C^1 function and

$$\|g_h'\|_\infty \leq \sqrt{\frac{2}{\pi}} K. \tag{3.3.11}$$

Proof. Since $0 \leq h \leq 1$ we deduce from the equality (3.3.6) coupled with the Mills ratio inequalities (1.1.2) that $g_h \in C_b^1(\mathbb{R})$

$$\sup_{x \in \mathbb{R}} \big| g_h'(x) \big| \leq \sqrt{\frac{2}{\pi}} K.$$

Clearly $\partial_x h \in L^2(\mathbb{R}, \boldsymbol{\Gamma}_1)$. We have $g_h'(x) = h(x) + x g_h(x) \in \mathbb{D}^{1,2}(\mathbb{R})$ so $g_h \in \mathbb{D}^{2,2}(\mathbb{R})$. From the equality $-\delta_x g_h = g'(x) - x g(x) = h$ and we conclude that

$$-\partial_x \delta_x g_h = \partial_x h \ \text{ a.e. on } \mathbb{R}.$$

Using the identity $[\partial_x, \delta_x] = \mathbb{1}$ we deduce $-\partial_x \delta_x = -\mathbb{1} - \delta_x \partial_x = (L - \mathbb{1})$. Thus

$$(L - \mathbb{1}) g_h = \partial_x h.$$

Since $g_h \in \mathbb{D}^{2,2}(\mathbb{R})$ we deduce

$$g_h = (L-1)^{-1} \partial_x h \stackrel{(3.2.19)}{=} -\int_0^\infty e^{-t} T_t[\partial_x h] dt.$$

Using Mehler's formula (3.2.16) we deduce

$$T_t[\partial_x h](x) = \int_{\mathbb{R}} h'(e^{-t}x + \sqrt{1-e^{-2t}} y) \boldsymbol{\Gamma}_1(dy).$$

We set $u_x := e^{-t}x + \sqrt{1-e^{-2t}} y$ and we observe that for fixed x we have

$$\frac{d}{dy} h(u_x) = h'(u_x) \frac{du_x}{dy} = \sqrt{1-e^{-2t}}\, h'(u_x) \Rightarrow h'(u_x) = \frac{1}{\sqrt{1-e^{-2t}}} \frac{d}{dy} h(u_x).$$

Hence

$$\begin{aligned} T_t[\partial_x h](x) &= \frac{1}{\sqrt{1-e^{-2t}}} \int_{\mathbb{R}} \frac{d}{dy} h(u_x) \boldsymbol{\Gamma}_1(dy) = \frac{1}{\sqrt{1-e^{-2t}}} \int_{\mathbb{R}} h(u_x) y \, \boldsymbol{\Gamma}_1(dy) \\ &= \frac{1}{\sqrt{1-e^{-2t}}} \mathbb{E}\big[h\big(e^{-t}x + \sqrt{1-e^{-2t}} N \big) N \big]. \end{aligned}$$

This proves (3.3.10).

Clearly g_h is a C^1-function. To prove the estimate (3.3.11), we derivate (3.3.10) we respect to x and we deduce

$$g_h'(x) = -\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbb{E}\big[\, h'\big(\, e^{-t}x + \sqrt{1-e^{-2t}}N\,\big) N\,\big]dt.$$

Since $|h'| \leq K$ we deduce

$$\big|\, g_h'(x)\,\big| \leq K\mathbb{E}[\,|N|\,]\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}dt = K\sqrt{\frac{2}{\pi}}\int_0^1 \frac{dv}{2\sqrt{1-v}} = K\sqrt{\frac{2}{\pi}}.$$

□

From the above proposition and the equality (3.3.9) we obtain immediately the following useful result.

Corollary 3.3.8. *Let $N \sim \mathcal{N}(0,1)$. Then for any square integrable random variable F we have*

$$\operatorname{dist}_{FM}(F,N) \leq \operatorname{dist}_W(F,N) \leq \sup_{g\in\mathcal{F}_W}\Big|\,\mathbb{E}[\, g'(F) - Fg(F)\,]\,\Big|, \tag{3.3.12}$$

where

$$\mathcal{F}_W := \left\{ g \in C^1(\mathbb{R});\ \ \|g'\|_\infty \leq \sqrt{\frac{2}{\pi}}\right\}. \tag{3.3.13}$$

□

3.3.3. The multidimensional Stein method. The Stein method has a multidimensional counterpart. To describe it we need to introduce some notation. Denote by $\operatorname{End}(\mathbb{R}^n)$ the space of linear operators $\mathbb{R}^n \to \mathbb{R}^n$. We define the *Hilbert-Schmidt* inner product on $\operatorname{End}(\mathbb{R}^n)$ to be

$$(A,B)_{HS} := \operatorname{tr} AB^* = \sum_{i,j} A_{ij}B_{ij},\ \ \forall A,B \in \mathcal{L}(\mathbb{R}^n).$$

The next result generalizes the one-dimensional Stein lemma

Lemma 3.3.9 (Multidimensional Stein lemma)**.** *Let $d \in \mathbb{N}$ and $C \in \operatorname{End}(\mathbb{R}^d)$ be a symmetric operator such that $C \geq 0$. Let $\boldsymbol{N} = (N_1,\dots,N_d)$ be a random d-dimensional vector. Then the following statements are equivalent.*

(i) $\boldsymbol{N} \sim \mathcal{N}(0,C)$

(ii) *For any C^2-function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded first and second order derivatives we have*

$$\mathbb{E}\big[\,\big(\boldsymbol{N}, \nabla f(\boldsymbol{N})\,\big)\,\big] = \mathbb{E}\big[\,\big(\, C, \operatorname{Hess} f(\boldsymbol{N})\,\big)_{HS}\,\big]. \tag{3.3.14}$$

Proof. (i) ⇒ (ii). If $C > 0$, then the implication follows from an immediate integration by parts and the equality

$$\boldsymbol{\Gamma}_C(dx) = \frac{1}{\sqrt{\det(2\pi C)}}e^{-\frac{1}{2}(C\boldsymbol{x},\boldsymbol{x})}d\boldsymbol{x}.$$

The general case follows from the general case applied to the nondegenerate matrices $C_\varepsilon = C + \varepsilon\mathbb{1}$ and then (carefully) letting $\varepsilon \to 0$.

(ii) ⇒ (i). Fix $\boldsymbol{G} \sim \mathcal{N}(0, C)$ independent of $\boldsymbol{N}$ and a C^2 function $f : \mathbb{R}^d \to \mathbb{R}$ as in (ii). We set

$$\varphi(t) := \mathbb{E}\big[\, f\big(\, \sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{G}\,\big)\,\big].$$

Then

$$\varphi(1) = \mathbb{E}[\, f(\boldsymbol{N})\,], \;\; \varphi(0) = \mathbb{E}[\, f(\boldsymbol{G})\,]$$

and thus

$$\mathbb{E}[\, f(\boldsymbol{N})\,] - \mathbb{E}[\, f(\boldsymbol{G})\,] = \int_0^1 \varphi'(t)dt$$

$$= \int_0^1 \mathbb{E}\big[\, \big(\, \nabla f(\, \sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{G}\,), \boldsymbol{N}\,\big)\,\big]\frac{dt}{2\sqrt{t}} - \int_0^1 \mathbb{E}\big[\, \big(\, \nabla f(\, \sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{G}\,), \boldsymbol{G}\,\big)\,\big]\frac{dt}{2\sqrt{1-t}}$$

Using (3.3.14) we deduce by conditioning on $\boldsymbol{G}$ that, for any $x \in \mathbb{R}^d$, we have

$$\mathbb{E}\big[\, \big(\, \nabla f(\, \sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{x}\,), \boldsymbol{N}\,\big)\,\big] = \underbrace{\sqrt{t}\mathbb{E}\big[\, \big(\, C, \operatorname{Hess} f(\sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{x}\,\big)_{HS}\,\big]}_{=:h_1(\boldsymbol{x},t)}.$$

Since $\boldsymbol{G} \sim \mathcal{N}(0, C)$ it satisfies (3.3.14) and, conditioning on $\boldsymbol{N}$, we deduce that for any $x \in \mathbb{R}^d$, we have

$$\mathbb{E}\big[\, \big(\, \nabla f(\, \sqrt{t}\boldsymbol{x} + \sqrt{1-t}\boldsymbol{G}\,), \boldsymbol{G}\,\big)\,\big] = \underbrace{\mathbb{E}\big[\, \big(\, C, \operatorname{Hess} f(\sqrt{t}\boldsymbol{x} + \sqrt{1-t}\boldsymbol{G})\,\big)_{HS}\,\big]}_{=:h_2(\boldsymbol{x},t)}.$$

Integrating $h_1(\boldsymbol{x}, t)$ and $h_2(\boldsymbol{x}, t)$ respectively with respect to the law of $\boldsymbol{G}$ and the law of $\boldsymbol{N}$, and then integrating with respect to t we deduce that

$$\boldsymbol{E}\big[\, f(\boldsymbol{N})\,\big] = \boldsymbol{E}\big[\, f(\boldsymbol{G})\,\big],$$

for any C^2-function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded first and second order derivatives. Since the class of such functions is separating we deduce that $\boldsymbol{N} \sim \boldsymbol{G} \sim \mathcal{N}(0, C)$. □

Definition 3.3.10. Let $\boldsymbol{N} \sim \mathcal{N}(0, \mathbb{1}_d)$ and $h : \mathbb{R}^n \to \mathbb{R}$ a measurable function such that $\mathbb{E}\big[\, |h(\boldsymbol{N})|\,\big] < \infty$. The *Stein's equation* associated to h and $\boldsymbol{N}$ is the p.d.e.

$$Lf(\boldsymbol{x}) = -\Delta f(\boldsymbol{x}) - \boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) = h(\boldsymbol{x}) - \mathbb{E}[\, h(\boldsymbol{N})\,], \;\; \Delta = \sum_{j=1}^d \partial_{x_j}^2. \tag{3.3.15}$$

Observe that if $h : \mathbb{R}^d \to \mathbb{R}$ is a Lipschitz continuous function, then the function

$$h_\perp(\boldsymbol{x}) := h(\boldsymbol{x}) - \mathbb{E}[\, h(\boldsymbol{N})\,] \in L^2(\mathbb{R}^d, \boldsymbol{\Gamma}) \;\text{ and }\; \int_{\mathbb{R}^d} h_\perp(\boldsymbol{x})\boldsymbol{\Gamma}(d\boldsymbol{x}) = 0.$$

Thus, $h_\perp$ lies in the range of the Ornstein-Uhlenbeck operator $L : \mathbb{D}^{2,2}(\mathbb{R}^d) \to \mathbb{D}^{0,2}(\mathbb{R}^d)$ so there exists a unique function $f_h \in \mathbb{D}^{2,2}(\mathbb{R}^d)$ such that

$$Lf_h(\boldsymbol{x}) = h_\perp(\boldsymbol{x}) \;\text{ and }\; \int_{\mathbb{R}^d} f_h(\boldsymbol{x})\boldsymbol{\Gamma}(d\boldsymbol{x}) = 0.$$

More precisely, $f_h = L^{-1}h = L^{-1}h_\perp$. We can now state the multidimensional counterpart of Proposition 3.3.7.

Proposition 3.3.11. *Let $h : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous function. Then the function*

$$f_h = L^{-1}h = L^{-1}h_\perp$$

is well defined, C^2 and admits the representation

$$f_h(\boldsymbol{x}) = -\int_0^\infty T_t[h_\perp]dt = \int_0^\infty \mathbb{E}\big[\, h(\boldsymbol{N}) - h\big(\, e^{-t}\boldsymbol{x} + \sqrt{1-e^{-2t}}\boldsymbol{N}\,\big)\,\big]dt. \tag{3.3.16}$$

Moreover, if $\mathrm{Lip}(h) \le K$ *then,*

$$\sup_{\boldsymbol{x}\in\mathbb{R}^d} \|\, \mathrm{Hess}\, f_h(\boldsymbol{x})\, \|_{HS} \le K\sqrt{d}. \tag{3.3.17}$$

Proof. Let $h_n \in L^2(\mathbb{R}^d, \boldsymbol{\Gamma})$ be the n-th chaos component of $h(x)$. Then, in L^2, we have the following equalities

$$h(x) = \sum_{n\ge 0} h_n(x), \quad h_\perp(x) = \sum_{n\ge 1} h_n(x),$$

$$L^{-1}h_\perp(x) = -\sum_{n\ge 1}\frac{1}{n}h_n(x) = -\sum_{n\ge 1}\int_0^\infty e^{-nt}h_n(x) = -\int_0^\infty T_t[h_\perp]dt.$$

This proves the first part of (3.3.16). The second part of this equality follows from Mehler's formula. The C^2-regularity of f_h is a consequence of basic elliptic regularity results.

To prove (3.3.17) we observe that

$$\partial^2_{x_i x_j} f_h(\boldsymbol{x}) = -\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbb{E}\Big[\,\partial_{x_j} h\big(\, e^{-t}\boldsymbol{x} + \sqrt{1-e^{-2t}}\boldsymbol{N}\,\big)N_i\,\Big]dt$$

Thus, if $B \in \mathrm{End}(\mathbb{R}^d)$, we have

$$\big|\,\big(\, B, \mathrm{Hess}\, f_h(\boldsymbol{x})\,\big)_{HS}\,\big| = \left|\sum_{i,j} \partial^2_{x_i x_j} f_h(\boldsymbol{x})\right|$$

$$= \left|\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbb{E}\Big[\,\Big(\, B\boldsymbol{N}, \nabla h\big(\, e^{-t}\boldsymbol{x} + \sqrt{1-e^{-2t}}\boldsymbol{N}\,\big)\Big)\,\Big]dt\right|dt$$

$$\le \|\nabla h\|_\infty \mathbb{E}\Big[\,\big|B\boldsymbol{N}\,\big|_{\mathbb{R}^d}\,\Big]\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}dt \le K\sqrt{d}\,\sqrt{\mathbb{E}\Big[\,\big|B\boldsymbol{N}\,\big|^2_{\mathbb{R}^d}\,\Big]},$$

because $\|\nabla h\|_\infty \le K\sqrt{d}$ and

$$\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}dt = 1.$$

A simple computation shows that

$$\mathbb{E}\Big[\,\big|B\boldsymbol{N}\,\big|^2_{\mathbb{R}^d}\,\Big] = \|B\|^2_{HS}.$$

This completes the proof of (3.3.17). □

Proposition 3.3.11 admits the following immediate generalization.

Proposition 3.3.12. *Fix a symmetric positive definite operator* $C \in \operatorname{End}(\mathbb{R}^d)$. *Denote by* $\lambda_{\min}(C)$ *and respectively* $\lambda_{\max}(C)$ *the smallest and the largest eigenvalue of* C. *Fix a random vector* $\boldsymbol{N} \sim \mathcal{N}(0, C)$ *and a Lipschitz continuous function* $h : \mathbb{R}^d \to \mathbb{R}$. *Set* $K := \operatorname{Lip}(h)$. *Then the function*

$$f_h(\boldsymbol{x}) = \int_0^\infty \mathbb{E}\big[\, h(\boldsymbol{N}) - h\big(\, e^{-t}\boldsymbol{x} + \sqrt{1 - e^{-2t}}\boldsymbol{N}\,\big)\,\big]\, dt \tag{3.3.18}$$

is well defined, it is C^2 *and satisfies the* Stein's equation

$$\Big(\, C, \operatorname{Hess} f(\boldsymbol{x})\,\Big)_{HS} - \big(\boldsymbol{x}, \nabla f(\boldsymbol{x})\,\big) = h(\boldsymbol{x}) - h(\boldsymbol{N}). \tag{3.3.19}$$

Moreover

$$\sup_{x \in \mathbb{R}^d} \big\|\operatorname{Hess} f_h(\boldsymbol{x})\,\big\|_{HS} \leq K \frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)}. \tag{3.3.20}$$

Main Idea. The above proposition can be obtained from Proposition 3.3.11 by choosing an orthonormal basis $\boldsymbol{f}_1, \dots, \boldsymbol{f}_d$ of $\mathbb{R}^d$ that diagonalizes C,

$$C\boldsymbol{f}_k = \lambda_k \boldsymbol{f}_k, \;\; k = 1, \dots, d, \;\; 0 < \lambda_1 \leq \cdots \leq \lambda_d.$$

□

The last result implies the following multi-dimensional counterpart of Corollary 3.3.8.

Corollary 3.3.13. *Fix a symmetric positive definite operator* $C \in \mathcal{L}(\mathbb{R}^d)$ *and a random vector* $\boldsymbol{N} \sim \mathcal{N}(0, C)$. *If* F *is a square integrable* $\mathbb{R}^d$*-valued random variable, then*

$$\operatorname{dist}_{FM}(F, \boldsymbol{N}) \leq \operatorname{dist}_W(F, \boldsymbol{N}) \leq \sup_{f \in \mathcal{F}_d} \Big|\, \mathbb{E}\Big[\, \big(\, C, \operatorname{Hess} f(F)\,\big)_{HS} - \big(F, \nabla f(F)\,\big)\,\Big]\,\Big|, \tag{3.3.21}$$

where $\mathcal{F}_d$ *consists of the* C^2*-functions* $f : \mathbb{R}^d \to \mathbb{R}$ *satisfying (3.3.20) with* $K = 1$. □

3.4. Wiener chaos limit theorems

The classical central limit theorem states that if $(X_n)_{n\geq 1}$ is a sequences of independent random variables, with mean zero and variance 1 then the random variables

$$F_\nu := \frac{1}{\sqrt{\nu}} \sum_{k=1}^{\nu} X_k$$

converge in distribution to a standard normal random variable. It classical proofs rely in an essential manner on the independence assumption. We will use the methods developed in the previous sections to prove central limit theorems involving sums of dependent random variables but such that long distance correlations are small, i.e.

$$\lim_{c \to infty} \sup_{n \in \mathbb{N}} \big|\, \operatorname{Cov}\big[\, X_n, X_{n+k}\,\big]\,\big| = 0.$$

The presentation is heavily inspired from the monograph [**124**]. For a continuously updated list of applications of this technique we refer to the webpage maintained by Ivan Nourdin

`https://sites.google.com/site/malliavinstein/home`

3.4.1. An abstract limit theorem. Fix a separable Gaussian Hilbert space $\mathfrak{X} \subset L^2(\Omega, \mathcal{S}, \mathbb{P})$. As usual, we set $L^2_{\mathfrak{X}}\big(\Omega\big) = \mathcal{F}(\mathfrak{X}) = L^2(\Omega, \mathcal{S}_{\mathfrak{X}}, \mathbb{P})$ and we denote by Proj_n the orthogonal projection onto the n-th chaos $\mathfrak{X}^{:n:}$. For any number $N \in \mathbb{N}_0$ we set

$$\mathrm{Proj}_{\leq N} = \bigoplus_{0 \leq n \leq N} \mathrm{Proj}_n, \;\; \mathrm{Proj}_{>N} = \mathbb{1} - \mathrm{Proj}_{\leq N}.$$

For $F \in L^2_{\mathfrak{X}}\big(\Omega\big)$ and $n \in \mathbb{N}_0$ we denote by $\mathrm{Var}_n(F)$ the variance of $\mathrm{Proj}_n(F)$. We have

$$\mathrm{Var}(F) = \sum_{n \geq 1} \mathrm{Var}_n(F),$$

and we set

$$\mathrm{Var}_{\leq N} := \sum_{n=1}^{N} \mathrm{Var}_n\big[\,F\,\big], \;\; \mathrm{Var}_{>N}\big[\,F\,\big] = \sum_{n>N} \mathrm{Var}_n(F) = \mathrm{Var}\big[\,F\,\big] - \mathrm{Var}_{\leq N}\big[\,F\,\big].$$

We begin by describing a simple sufficient condition guaranteeing the convergence in law to a normal random variable of a sequence of random variables in $L^2_{\mathfrak{X}}\big(\Omega\big)$.

Proposition 3.4.1. *Consider a sequence of random variables* $(F_\nu)_{\nu \geq 1}$ *in* $L^2_{\mathfrak{X}}\big(\Omega\big)$ *such that*

$$\mathbb{E}\big[\,F_\nu\,\big] = 0, \;\; \forall \nu,$$

i.e., $\mathrm{Proj}_0(F_\nu) = 0$, $\forall \nu$. *Suppose that the following hold.*

(C_1) *For any* $n \in \mathbb{N}$, *the sequence of variances* $\mathrm{Var}_n\big[\,F_\nu\,\big]$ *converges as* $\nu \to \infty$ *to a nonnegative number* v_n.

(C_2) *The sequence*

$$V_N := \sup_{\nu \geq 1} \mathrm{Var}_{>N}\big[\,F_\nu\,\big]$$

converges to 0 *as* $N \to \infty$. *In other words, as* $N \to \infty$, *the "tails"* $\mathrm{Proj}_{>N} F_\nu$ *converge to* 0 *in* L^2, uniformly *with respect to* ν.

(C_3) *For any* $N > 0$ *the sequence of random variables* $\mathrm{Proj}_{\leq N}(F_\nu)$ *converges in law to a normal random variable.*

Then the following hold.

(i) *The series* $\sum_{n \geq 1} v_\nu$ *is convergent. We denote by* v *its sum.*

(ii)

$$\lim_{\nu \to \infty} \mathrm{Var}\big[\,F_\nu\,\big] = v.$$

(iii) *As* $\nu \to \infty$, *the random variable* F_ν *converges in law to a random variable* $F_\infty \sim \mathcal{N}(0, v)$.

Proof. (i) Fix $\varepsilon > 0$. We can find $N(\varepsilon) > 0$ such that for any $N > N(\varepsilon)$ we have $V_N < \varepsilon$. For all $n > m > N(\varepsilon)$ we have

$$\sum_{k=m}^{n} \mathrm{Var}_k\big[\,F_\nu\,\big] \leq \sum_{k>N} \mathrm{Var}_k\big[\,F_\nu\,\big] \leq V_N < \varepsilon$$

which shows that

$$\forall n > M > N(\varepsilon): \;\; \sum_{k=m}^{n} v_k = \lim_{\nu \to \infty} \sum_{k=m}^{n} \mathrm{Var}_k\big[\,F_\nu\,\big] \leq \varepsilon.$$

To prove (ii) observe that for any $N > 0$ we have

$$\Big|\operatorname{Var}\big[\,F_\nu\,\big] - v\big| \leq \sum_{n\leq N}\Big|\,|\operatorname{Var}_n\big[\,F_\nu\,\big] - v_n\Big| + \sum_{n>N}\operatorname{Var}_n(F_\nu) + \sum_{n>N} v_n$$

$$\leq \sum_{n\leq N}\Big|\operatorname{Var}_n(F_\nu) - v_n\Big| + V_N + \sum_{n>N} v_n$$

This proves that

$$\limsup_{\nu\to\infty}\big|\operatorname{Var}\big[\,F_\nu\,\big] - v\,\big] \leq V_N + \sum_{n>N} v_n,\ \ \forall N > 0.$$

The conclusion (ii) is obtained by letting $N \to \infty$ in the above inequality.

(iii) Let $X \in \mathfrak{X}$, $\|X\| = 1$, so that $X \in \mathcal{N}(0,1)$, $\sqrt{v}X \in \mathcal{N}(0,v)$. We will show that for any bounded Lipschitz function $h : \mathbb{R} \to \mathbb{R}$ we have

$$\lim_{\nu\to\infty}\mathbb{E}\big[\,h(F_\nu)\,\big] = \mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]. \tag{3.4.1}$$

Observe that if $v = 0$, we deduce from (ii) that $F_\nu \to 0$ in L^2 so F_ν converges in law to the degenerate normal random variable with variance 0. Assume $v > 0$. Without loss of generality we can assume $v = 1$.

Fix a bounded Lipschitz function $h : \mathbb{R} \to \mathbb{R}$ and set

$$K := \|h\|_\infty + \operatorname{Lip}(h).$$

For $N > 0$ we set

$$G_{\nu,N} = \operatorname{Proj}_{\leq N}(F_\nu),\ \ H_{\nu,N} = F_\nu - G_{\nu,N} = \operatorname{Proj}_{>N}(F_\nu)$$

$$v_N = \sum_{n\leq N} v_n,\ \ \sigma_N = \sqrt{v_N}$$

so that, as $\nu \to \infty$ $G_{\nu,N}$ converges in law to $\sigma_N X$ and $H_{\nu,N}$ converges in L^2 to 0.

$$\Big|\mathbb{E}\big[\,h(F_\nu)\,\big] - \mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]\,\Big| \leq \Big|\mathbb{E}\big[\,h(F_\nu)\,\big] - \mathbb{E}\big[\,h(G_{\nu,N})\,\big]\,\Big|$$

$$+\Big|\mathbb{E}\big[\,h(G_{\nu,N})\,\big] - \mathbb{E}\big[\,h(\sigma_N X)\,\big]\,\Big| + \Big|\mathbb{E}\big[\,h(\sigma_N X)\,\big] - \mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big].\,\Big|$$

Now observe that

$$\Big|\mathbb{E}\big[\,\big|\,h(G_{\nu,N} + H_{\nu,N})\ -\ h(G_{\nu,N})\,\big|\,\big]\,\Big| \leq K\mathbb{E}\big[\,\big|H_{\nu,N}\,\big|\,\big] \leq K\big\|H_{\nu,N}\|_{L^2},$$

$$\lim_{\nu\infty}\big\|H_{\nu,N}\|_{L^2} = 0,\ \ \lim_{\nu\to\infty}\Big|\mathbb{E}\big[\,h(G_{\nu,N})\,\big] - \mathbb{E}\big[\,h(\sigma_N X)\,\big]\,\Big| = 0,$$

so that

$$\limsup_{\nu\to\infty}\Big|\,\mathbb{E}\big[\,h(F_\nu)\,\big] - \mathbb{E}\big[\,h(X)\,\big]\,\Big| \leq \Big|\,\mathbb{E}\big[\,h(\sigma_N X)\,\big] - \mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]\,\Big|.$$

Letting $N \to \infty$ we deduce

$$\lim_{\nu\to\infty}\Big|\,\mathbb{E}\big[\,h(F_\nu)\,\big] - \mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]\,\Big| = 0,$$

for any bounded Lipschitz function h. This proves (iii).

□

In the remainder of this section we will explain how to combine the Stein method with the Malliavin calculus to prove central limit results of the type described in Proposition 3.4.1, with condition C_3 replaced by one that is easier to verify in concrete situations. These techniques were pioneered by D. Nualart and G. Peccati in [**127**] and have since generated a lot of follow-up investigations[8]; see e.g. [**123, 125**] and the references therein. We follow the presentation in the award winning monograph of I. Nourdin and G. Peccati, [**124**].

3.4.2. Central limit theorem: single chaos. The following proposition is the key result in the implementation of the Stein method in the Wiener chaos context.

Proposition 3.4.2 (Key abstract estimate)**.** *Let $F \in \mathbb{D}^{1,2}(\mathfrak{X})$ such that*

$$\mathbb{E}\big[\,F\,\big] = 0,\ \ \mathbb{E}\big[\,F^2\,\big] = 1.$$

If $g : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function and $K = \operatorname{Lip}(f)$, then

$$\boxed{\big|\,\mathbb{E}\big[\,g'(F)\,\big] - \mathbb{E}\big[\,Fg(F)\,\big]\,\big| \leq K \cdot \Big|\,\mathbb{E}\Big[\,\Big(1 - \big(\,DF, -DL^{-1}F\big)_{\mathfrak{X}}\,\Big)\,\Big]\,\Big|}. \tag{3.4.2}$$

Proof. Note first that g' is defined only a.e.. However, according to Theorem 3.2.10, F is nonconstant so the law of F has a density, and thus the random variable $g'(F)$ is a.s. well defined. Using the integration-by-parts formula (3.2.20) with $F = G$ we deduce

$$\begin{aligned}\big|\,\mathbb{E}\big[\,g'(F)\,\big] - \mathbb{E}\big[\,Fg(F)\,\big]\,\big| &= \Big|\,\mathbb{E}\Big[\,g'(F)\Big(\,1 - \big(\,DF, -DL^{-1}F\big)_{\mathfrak{X}}\,\Big)\,\Big]\,\Big| \\ &\leq K \cdot \Big|\,\mathbb{E}\Big[\,\Big(1 - \big(\,DF, -DL^{-1}F\big)_{\mathfrak{X}}\,\Big)\,\Big]\,\Big|.\end{aligned}$$

□

Corollary 3.4.3. *Let $F \in \mathbb{D}^{1,2}(\mathfrak{X})$ with $\mathbb{E}[F] = 0$, $\mathbb{E}[F^2] = \sigma^2 > 0$. If $N \sim \mathcal{N}(0, \sigma^2)$, then*

$$\operatorname{dist}_W(F, N) \leq \sqrt{\frac{2}{\pi\sigma^2}}\,\mathbb{E}\big[\,\big|\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\,\big|\,\big]. \tag{3.4.3}$$

If, in addition, $F \in \mathbb{D}^{1,4}(\mathfrak{X})$, then

$$\mathbb{E}\Big[\,\Big|\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\,\Big|\,\Big] \leq \sqrt{\operatorname{Var}\Big[\,\big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\Big]}. \tag{3.4.4}$$

Proof. The case $\sigma = 1$ follows from Corollary 3.3.8 and the inequality (3.4.2). The general case of (3.4.3) follows from the case $\sigma = 1$ applied to the new random variable $\sigma^{-1}F$.

To prove (3.4.4) we observe that

$$\mathbb{E}\Big[\,\Big|\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\,\Big|\,\Big] \leq \sqrt{\mathbb{E}\big[\,\big(\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\,\big)^2\,\big]}.$$

From the integration by parts formula (3.2.20) we deduce that

$$\mathbb{E}\Big[\,\big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\,\Big] = \sigma^2,$$

so that,

$$\mathbb{E}\Big[\,\big(\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\,\big)^2\,\Big] = \operatorname{Var}\Big[\,\big(DF, -DL^{-1}F\,\big)_{\mathfrak{X}}\Big].$$

[8] Ivan Nourdin maintains a site dedicated to this novel way of approaching limit theorems https://sites.google.com/site/malliavinstein/home.

To show that the above variance is finite observe that

$$\mathbb{E}\Big[\,\big(DF,-DL^{-1}F\,\big)^2_{\mathfrak{X}}\,\Big]\leq\sqrt{\mathbb{E}\Big[\,\|DF\|^4_{\mathfrak{X}}\,\Big]}\cdot\sqrt{\mathbb{E}\Big[\,\|DL^{-1}F\|^4_{\mathfrak{X}}\,\Big]}.$$

The Kree-Meyer inequalities (3.2.27) imply that the quantities in the right-hand-side above are finite. □

Remark 3.4.4. The method of proof of Proposition 3.4.2 and the statement of Corollary 3.4.3 rely on the assumption $\sigma>0$ which may not be easy to verify in some concrete situations. □

Proposition 3.4.5. *Let $F\in\mathbb{D}^{1,2}(\mathfrak{X})$ such that $\mathbb{E}[F]=0$, $\mathbb{E}[F^2]=\sigma^2$. If $h\in C^2_b(\mathbb{R})$ and $N\sim\mathcal{N}(0,\sigma^2)$, then*

$$\big|\,\mathbb{E}\big[\,h(F)\,\big]-\mathbb{E}\big[\,h(N)\,\big]\,\big|\leq\frac{1}{2}\|h''\|_\infty\cdot\mathbb{E}\Big[\,\Big|\,\big(\,DF,-DL^{-1}F\big)_{\mathfrak{X}}\,\big)-\sigma^2\,\Big|\,\Big].\tag{3.4.5}$$

In particular, if $F\in\mathbb{D}^{1,4}$, then

$$\operatorname{dist}_{C^2}(F,N)\leq\frac{1}{2}\mathbb{E}\Big[\,\Big|\,\big[\,DF,-DL^{-1}F\,\big]_{\mathfrak{X}}\,\big)-\sigma^2\,\Big|\,\Big]\leq\frac{1}{2}\sqrt{\operatorname{Var}\Big[\,\big[\,DF,-DL^{-1}F\,\big]_{\mathfrak{X}}\,\Big]}.\tag{3.4.6}$$

Proof. The results is obviously true if $\sigma^2=0$ so we can assume that $\sigma^2>0$. We set

$$\varphi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbb{E}\big[\,h\big(\,e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\,\big)\,\big]dx.$$

Note that

$$\varphi(\infty)=\mathbb{E}\big[\,h(F)\,\big],\;\;\varphi(0)=\mathbb{E}\big[\,h(N)\,\big],$$

so that

$$\mathbb{E}\big[\,h(F)\,\big]-\mathbb{E}\big[\,h(N)\,\big]=\int_0^\infty\varphi'(t)dt.$$

We have

$$\varphi'(t)=\frac{e^{-t}\sigma}{\sqrt{2\pi}}\mathbb{E}\left[\int_{-\infty}^{\infty}h'\big(\,e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\,\big)xe^{-\frac{x^2}{2}}dx\right]$$
$$+\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\times\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F\mathbb{E}\Big[\,h'\big(\,e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\,\big)\,\Big]e^{-\frac{x^2}{2}}dx.$$

Performing an usual integration by parts in the first integral and using the Malliavin integration by parts formula (3.2.20) in the second integrand we deduce

$$\varphi'(t)=-\frac{e^{-2t}\sigma^2}{\sqrt{2\pi}}\mathbb{E}\left[\int_{-\infty}^{\infty}h''\big(\,e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\,\big)e^{-\frac{x^2}{2}}dx\right]$$
$$+\frac{e^{-2t}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbb{E}\Big[\,h''\big(\,e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\,\big)\big(DF,-DL^{-1}F\,\big)_{\mathfrak{X}}\,\Big]e^{-\frac{x^2}{2}}dx$$
$$=\frac{e^{-2t}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbb{E}\Big[\,h''\Big(\,e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\,\Big)\cdot\Big(\,\big(DF,-DL^{-1}F\,\big)_{\mathfrak{X}}-\sigma^2\,\Big)\,\Big]e^{-\frac{x^2}{2}}dx.$$

We deduce

$$\mathbb{E}\big[\,h(F)\,\big]-\mathbb{E}\big[\,h(N)\,\big]$$
$$=\int_{-\infty}^{\infty}\frac{e^{-2t}dt}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbb{E}\Big[\,h''\Big(\,e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\,\Big)\cdot\Big(\,\big(DF,-DL^{-1}F\,\big)_{\mathfrak{X}}-\sigma^2\,\Big)\,\Big]e^{-\frac{x^2}{2}}dx.$$

We reach the desired conclusion by observing that

$$\mathbb{E}\Big[\Big|h''\Big(e^{-t}\sigma x+\sqrt{1-e^{-2t}}F\Big)\Big|\Big]\leq\|h''\|_\infty,\ \ \forall x.$$

□

Observe that when $F\in\mathfrak{X}^{:q:}$, $q>0$, then $F\in\mathbb{D}^{1,4}$ and

$$\big(DF,-DL^{-1}F\big)_{\mathfrak{X}}=\frac{1}{q}\|DF\|^2_{\mathfrak{X}}.$$

In this case we can provide more detailed information. This will require a bit of Itô calculus and a bit more terminology.

Given $p,q\in\mathbb{N}$ and $r\in\mathbb{N}_0$ such that $r\leq\min\{p,q\}$ we define the map

$$\otimes_r:\mathfrak{X}^{\otimes p}\times\mathfrak{X}^{\otimes q}\to\mathfrak{X}^{\otimes(p+q-2r)}$$

to be the unique continuous bilinear map such that

$$(X_1\otimes\cdots\otimes X_p)\otimes_r(Y_1\otimes\cdots\otimes Y_q)=\left(\prod_{j=1}^r\mathbb{E}\big[\,X_jY_j\,\big]\right)X_{r+1}\otimes\cdots\otimes X_p\otimes Y_{r+1}\otimes\cdots\otimes Y_q.$$

This induces a map

$$\tilde{\boxtimes}_r:\mathfrak{X}^{\hat{\odot}p}\times\mathfrak{X}^{\hat{\odot}q}\to\mathfrak{X}^{\odot(p+q-2r)}$$

to be

$$u\tilde{\boxtimes}_r v:=\mathbf{Sym}\big[\,u\otimes_r v\,\big],\ \ \forall u\in\mathfrak{X}^{\odot p},\ \ v\in\mathfrak{X}^{\odot q}.$$

Remark 3.4.6. If $W:L^2\big(T,\mathcal{M},\mu\big)\to\mathfrak{X}$ is a white noise isomorphism, $(T,\mathcal{M},\mu)$ convenient probability space, then we can *isometrically* identify $\mathfrak{X}^{\otimes p}$ with the space $L^2\big(T^p,\mathcal{M}^{\otimes p},\mu^{\otimes p}\big)$. Thus we can view $f\in\mathfrak{X}^{\otimes p}$ and $g\in\mathfrak{X}^{\otimes q}$ as L^2-functions

$$f:T^p\to\mathbb{R},\ \ g:T^q\to\mathbb{R}.$$

Then $f\otimes_r g$ can be identified with the function

$$f\boxtimes_r g:T^{p-r}\times T^{q-r}\to\mathbb{R}$$

given by

$$f\boxtimes_r g(x_{r+1},\dots,x_p,y_{r+1},\dots,y_q)$$
$$=\int_{T^r}f(t_1,\dots,t_r,x_{r+1},\dots,x_p)g(t_1,\dots,t_r,y_{r+1},\dots y_q)\mu^{\otimes r}\big[\,dt_1\cdots dt_r\,\big].$$

We set

$$f\tilde{\boxtimes}_r g=\mathbf{Sym}\,\big[\,f\boxtimes_r g\,\big].$$

Recall that $H^{\hat{\odot}q}$ denotes the q-the symmetric tensor power of H.

Lemma 3.4.7. *Let $q\in\mathbb{N}$, $q\geq 2$ and $f\in L^2\big(T,\mathcal{M},\mu\big)^{\hat{\odot}q}$. Set $F=\mathfrak{I}_p[f]$. Then the following hold.*

$$\frac{1}{q}\|\,DF\|^2_{\mathfrak{X}}=\mathbb{E}[F^2]+q\sum_{r=1}^{q-1}(r-1)!\binom{q-1}{r-1}^2\mathfrak{I}_{2q-2r}[f\tilde{\boxtimes}_r f],\tag{3.4.7a}$$

$$\mathrm{Var}\left(\frac{1}{q}\|\,DF\|^2_{\mathfrak{X}}\right)=\frac{1}{q^2}\sum_{r=1}^{q-1}r^2(r!)^2\binom{q}{r}^2(2q-2r)!\|f\tilde{\boxtimes}_r f\|^2_{\mathfrak{X}^{\odot(2q-2r)}},\tag{3.4.7b}$$

$$\boxed{\mathbb{E}\big[\,F^4\,\big]-3\mathbb{E}\big[\,F^2\,\big]^2}=\frac{3}{q}\sum_{r=1}^{q-1}r^2(r!)^2\binom{q}{r}^2(2q-2r)!\,\|\,f\,\tilde{\boxtimes}_r f\,\|^2_{\mathfrak{X}^{\otimes(2q-2r)}}$$
$$=\sum_{r=1}^{q-1}(q!)^2\binom{q}{r}^2\left(\|\,f\,\boxtimes_r f\|^2_{\mathfrak{X}^{\otimes(2q-2r)}}+\binom{2q-2r}{q-r}\|\,f\,\tilde{\boxtimes}_r f\,\|^2_{\mathfrak{X}^{\otimes(2q-2r)}}\right), \tag{3.4.7c}$$

$$\operatorname{Var}\left(\frac{1}{q}\|\,DF\|^2_{\mathfrak{X}}\right)\leq\frac{q-1}{3q}\boxed{\left(\mathbb{E}\big[\,F^4\,\big]-3\mathbb{E}\big[\,F^2\,\big]^2\right)}\leq(q-1)\operatorname{Var}\left(\frac{1}{q}\|\,DF\|^2_{\mathfrak{X}}\right). \tag{3.4.7d}$$

About the proof. Let us point out that (3.4.7b) follows immediately from (3.4.7a) via the isometry (3.1.41). The inequality (3.4.7d) follows immediately from (3.4.7b, 3.4.7c). Thus it suffices to prove only (3.4.7a) and (3.4.7c).

To prove (3.4.7a) it is convenient to consider a more general problem, that of finding the chaos decomposition of

$$\big[\,DF,DG\,\big]_{\mathfrak{X}},\quad F,G\in\mathfrak{X}^{:q:}$$

We write $F=\mathfrak{I}_p[f]$, $G=\mathfrak{I}_p[g]$, $f,g\in\mathfrak{X}^{\odot q}$. Using the polarization trick we can reduce the problem to the special case

$$f=X^{\otimes q},\;\; g=Y^{\otimes q},\;\; X,Y\in\mathfrak{X},\;\;\mathbb{E}\big[\,X^2\,\big]=\mathbb{E}\big[\,Y^2\,\big]=1.$$

Thus

$$F=H_q(X),\;\; DF=qH_{q-1}(X)DX,$$
$$G=H_q(Y),\;\; DG=qH_{q-1}(Y)DY,$$
$$\big(\,DF,DG\,\big)_{\mathfrak{X}}=q^2H_{q-1}(X)H_{q-1}(Y)\mathbb{E}[XY].$$

The equality (3.4.7a) now follows by invoking (3.1.29), (3.1.30) and the isometry equality (3.2.13).

The proof of (3.4.7c) requires a bit more work. The hardest part is the 2nd half of this equality. It is based on the (non-obvious) elementary identity

$$\boxed{(2q)!\|f\tilde{\boxtimes}f\|^2_{\mathfrak{X}^{\otimes(2q)}}=2(q!)^2\|f\|^4_{\mathfrak{X}^{\otimes q}}+(q!)^2\sum_{r=1}^{q-1}\binom{q}{r}^2\|f\otimes_r f\|^2_{\mathfrak{X}^{\boxtimes(2q-2r)}},\quad f\in\mathfrak{X}^{\odot q}}. \tag{3.4.8}$$

A convenient way to prove this is to use a white noise isomorphism as in Remark 3.4.6. We refer to [**124**, Lemma 5.2.4] for details. □

Corollary 3.4.8 (The fourth moment theorem, [**127**])**.** *Suppose that $F\in\mathfrak{X}^{:q:}$, $q\geq2$ and $\mathbb{E}[F^2]=\sigma^2>0$. Then for $N\in\mathcal{N}(0,\sigma)$ we have*

$$\operatorname{dist}_W(F,N)\leq\frac{1}{\sigma}\sqrt{\operatorname{Var}\Big(\frac{2}{q\pi}\|DF\|^2_{\mathfrak{X}}\Big)}\leq\frac{1}{\sigma}\sqrt{\frac{(2q-2)\big(\mathbb{E}\big[\,F^4\,\big]-3\sigma^4\big)}{3\pi q}}. \tag{3.4.9}$$

Thus, given a sequence $(F_n)_{n\geq0}$ in $\mathfrak{X}^{:q:}$, $q\geq2$ and $N\sim\mathcal{N}(0,\sigma)$ the following statements are equivalent.

(i) *The sequence $(F_n)_{n\geq0}$ converges in probability to N.*

(ii) *As $n\to\infty$, $\mathbb{E}\big[\,F_n^2\,\big]\to\mathbb{E}\big[\,N^2\,\big]=\sigma^2$ and $\mathbb{E}\big[\,F_n^4\,\big]\to\mathbb{E}\big[\,N^4\,\big]=3\sigma^4$.*

(iii) *If $F_n=\mathfrak{I}_q[f_n]$, $f_n\in\mathfrak{X}^{\odot q}$, then*

$$\lim_{n\to\infty}\|f_n\tilde{\boxtimes}_r f_n\|_{\mathfrak{X}^{\odot(2q-2r)}}=0,\;\;\forall r=1\dots,q-1.$$

(iv) $\operatorname{Var}\big(\,\|DF_n\|^2\,\big) \to 0$ *as* $n \to \infty$.

Proof. In this case we have

$$\big(\,DF, -DL^{-1}F\,\big)_{\mathfrak{X}} = \frac{1}{q}\big\|\,DF\,\big\|^2_{\mathfrak{X}}.$$

The desired conclusions follow from Corollary 3.4.3, (3.4.7b) and (3.4.7d). □

3.4.3. Central limit theorem: multiple chaoses. The results proved in the previous subsection have a multidimensional counterpart. The next result, is the multi-dimensional counterpart of Proposition 3.4.2 and Corollary 3.4.3

Proposition 3.4.9. *Fix* $d \geq 2$ *and let* $\boldsymbol{F} = (F_1, \dots, F_d)$ *be a random vector such that* $F_1, \dots, F_d \in \mathbb{D}^{1,4}(\mathfrak{X})$ *with* $\mathbb{E}\big[\,F_i\,\big] = 0$, i. *Let* $C \in \mathcal{L}(\mathbb{R}^d)$ *be a symmetric* positive *definite operator and let* $\boldsymbol{N} \sim \mathcal{N}(0, C)$. *Then*

$$\operatorname{dist}_W(\boldsymbol{F}, \boldsymbol{N}) \leq \frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)}\sqrt{\sum_{i,j=1}^d \mathbb{E}\Big[\,\big(\,C_{ij} - \big(\,DF_i, -DL^{-1}F_j\big)_{\mathfrak{X}}\,\big)^2\,\Big]} \tag{3.4.10}$$

Proof. Let M be the random operator $M : \Omega \to \mathcal{L}(\mathbb{R}^d)$ with the (i,j)-th entry given by

$$M_{ij} := \big(\,DF_j, -DL^{-1}F_i\big)_{\mathfrak{X}}.$$

Arguing as in the proof of Corollary 3.4.3 we deduce that $M_{ij} \in L^2$ since $F_i, F_j \in \mathbb{D}^{1,4}(\mathfrak{X})$. For $g \in C^2(\mathbb{R}^d)$ such that

$$\sup_{x \in \mathbb{R}^d}\big\|\,\operatorname{Hess} g(x)\,\big\|_{HS} \leq \frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)}$$

we have

$$\Big|\,\mathbb{E}\big[\,(C, \operatorname{Hess} g)_{HS}(\boldsymbol{F}) - (\boldsymbol{F}, \nabla g(\boldsymbol{F}))_{\mathbb{R}^d}\,\big]\,\Big| = \left|\sum_{i,j=1}^d C_{ij}\mathbb{E}\big[\,\partial^2_{x_i x_j} g(\boldsymbol{F})\,\big] - \sum_{i=1}^d \mathbb{E}\big[\,F_i \partial_{x_i} g(\boldsymbol{F})\,\big]\right|$$

(use the integration by parts formula (3.2.20))

$$= \left|\sum_{i,j=1}^d C_{ij}\mathbb{E}\big[\,\partial^2_{x_i x_j} g(\boldsymbol{F})\,\big] - \sum_{i,j=1}^d \mathbb{E}\big[\,\partial^2_{x_i x_j} g(\boldsymbol{F})(DF_j, -DL^{-1}F_i)_{\mathfrak{X}}\,\big]\right|$$

$$= \left|\sum_{i,j=1}^d C_{ij}\mathbb{E}\big[\,\partial^2_{x_i x_j} g(F)\big(\,C_{ij} - (DF_j, -DL^{-1}F_i)_{\mathfrak{X}}\,\big]\right|$$

$$= \Big|\,\mathbb{E}\big[\,(\operatorname{Hess} g(\boldsymbol{F}), C - M)_{HS}\,\big]\,\Big| \leq \sqrt{\mathbb{E}\big[\,\|\operatorname{Hess} g(\boldsymbol{F})\|^2_{HS}\,\big]} \cdot \sqrt{\mathbb{E}\big[\,\|C - M\|^2_{HS}\,\big]}$$

$$\leq \frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)}\sqrt{\mathbb{E}\big[\,\|C - M\|^2_{HS}\,\big]}.$$

We conclude by invoking Corollary 3.3.13. □

The next result, is the multidimensional counterpart of Proposition 3.4.5 and explains what to do when the covariance matrix C is possible degenerate.

Proposition 3.4.10. *Fix $d \geq 2$ and let $\boldsymbol{F} = (F_1, \ldots, F_d)$ be a random vector such that $F_1, \ldots, F_d \in \mathbb{D}^{1,4}(\mathfrak{X})$ with $\mathbb{E}[F_i] = 0$, i. Let $C \in \mathcal{L}(\mathbb{R}^d)$ be a symmetric, nonnegative definite operator and let $\boldsymbol{N} \sim \mathcal{N}(0, C)$. Then for every $h \in C^2(\mathbb{R}^d)$ such that $\|h''\|_\infty < \infty$ we have*

$$\big| \, \mathbb{E}\big[\, h(\boldsymbol{F})\,\big] - \mathbb{E}\big[\, h(\boldsymbol{N})\,\big] \,\big| \leq \frac{1}{2}\|h''\|_\infty \sqrt{\sum_{i,j=1}^{d} \mathbb{E}\Big[\, \big(\, C_{ij} - \big(\, DF_j, -DL^{-1}F_i\big)_{\mathfrak{X}}\,\big)^2\,\Big]} \tag{3.4.11}$$

Proof. Without any loss of generality we can assume $\boldsymbol{N}$ is independent of the Gaussian space $\mathfrak{X}$. Let h as in the statement of the proposition. For $t \in [0,1]$ we set

$$\Psi(t) := \mathbb{E}\big[\, h(\sqrt{1-t}\,\boldsymbol{F} + \sqrt{t}\,\boldsymbol{N})\,\big].$$

Then

$$\boldsymbol{E}\big[\, h(\boldsymbol{N})\,\big] - \boldsymbol{E}\big[\, h(\boldsymbol{F})\,\big] = \Psi(1) - \Psi(0) = \int_0^1 \Psi'(t)dt.$$

We have

$$\Psi'(t) = \sum_{i=1}^{d} \mathbb{E}\left[\, \partial_{x_i} h(\sqrt{1-t}\,\boldsymbol{F} + \sqrt{t}\,\boldsymbol{N})\, \left(\frac{1}{2\sqrt{t}} N_i - \frac{1}{2\sqrt{1-t}} F_i\right)\,\right].$$

At this point we want to use the following elementary but useful identity.

Lemma 3.4.11. *If $f = f(y_1, \ldots, y_d) : \mathbb{R}^d \to \mathbb{R}$ is C^1 with bounded derivatives, $\widehat{N} \sim \mathcal{N}(0, \mathbb{1}_d)$ and $T, S \in \mathcal{L}(\mathbb{R}^d)$, then*

$$\mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big)\big(\, T\hat{\boldsymbol{N}}\,\big)_i\,\big] = \sum_{k=1}^{d} \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big)(TS^*)_{ik}\,\big], \tag{3.4.12}$$

where $\big(\, T\hat{\boldsymbol{N}}\,\big)_i$ denotes the i-th component of the random vector $T\hat{\boldsymbol{N}}$ and $(TS^*)_{ik}$ denote the (i,k)-entry of the matrix TS^*

Proof of the lemma. We have

$$\mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big)\big(\, T\hat{\boldsymbol{N}}\,\big)_i\,\big] = \sum_{j=1}^{d} \mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big) T_{ij} N_j\big]$$

$(\delta_j = -\partial_{N_j} + N_j)$

$$= \sum_{j=1}^{d} \mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big) T_{ij} \delta_j(1)\big]$$

(integrate by parts using the equalities $\partial_{N_j} = \sum_k \partial_{y_k} \partial_{N_j} y_k$, $y_k = \sum_j S_{kj} N_j$)

$$= \sum_{j=1}^{d} \sum_{k=1}^{d} \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big) S_{kj} T_{ij}\big] = \sum_{k=1}^{d} \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big)(TS^*)_{ik}\,\big].$$

□

Now observe that if $f : \mathbb{R}^d \to \mathbb{R}$ is a C^1-function with bounded derivatives, and $\hat{\boldsymbol{N}} \sim \mathcal{N}(0, \mathbb{1}_d)$ is such that, $\boldsymbol{N} = \sqrt{C}\hat{\boldsymbol{N}}$, then (3.4.12) shows that

$$\mathbb{E}\big[\, f(\boldsymbol{N}) N_i\,\big] = \mathbb{E}\big[\, f(\, \sqrt{C}\hat{\boldsymbol{N}}\,)\big(\sqrt{C}\hat{\boldsymbol{N}}\big)_i\,\big] = \sum_{k=1}^{d} \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big) C_{ik}\,\big]. \tag{3.4.13}$$

We have

$$\mathbb{E}\Big[\,\partial_{x_i}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N})N_i\Big]=\mathbb{E}_{\boldsymbol{x}}\Big[\,\mathbb{E}\Big[\,\partial_{x_i}h(\sqrt{1-t}\,\boldsymbol{x}+\sqrt{t}\,\boldsymbol{N})N_i\Big|\boldsymbol{F}=\boldsymbol{x}\Big]\,\Big]$$

$$\stackrel{(3.4.13)}{=}\sqrt{t}\sum_{j=1}^{d}\mathbb{E}_{\boldsymbol{x}}\Big[\,C_{ij}\mathbb{E}\Big[\,\partial^2_{x_ix_j}h(\sqrt{1-t}\,\boldsymbol{x}+\sqrt{t}\,\boldsymbol{N})\Big|\boldsymbol{F}=\boldsymbol{x}\Big]\,\Big]$$

$$=\sqrt{t}\sum_{j}C_{ij}\mathbb{E}\Big[\,\partial^2_{x_ix_j}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N})\,\Big].$$

Using the integration by parts formula (3.2.20) we deduce

$$\mathbb{E}\Big[\,\partial_{x_i}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N})F_i\Big]=\mathbb{E}_{\boldsymbol{x}}\Big[\,\mathbb{E}\Big[\,\partial_{x_i}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{x})F_i\Big|\boldsymbol{N}=\boldsymbol{x}\Big]\,\Big]$$

$$=\sqrt{1-t}\sum_{j=1}^{d}\mathbb{E}_{\boldsymbol{x}}\Big[\,\mathbb{E}\Big[\,\partial^2_{x_ix_j}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{x})\big(DF_j,-DL^{-1}F_i\big)_{\mathfrak{X}}\Big|\boldsymbol{N}=\boldsymbol{x}\Big]\,\Big]$$

$$=\sqrt{1-t}\sum_{j=1}^{d}\mathbb{E}\Big[\,\partial^2_{x_ix_j}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N})\big(DF_j,-DL^{-1}F_i\big)_{\mathfrak{X}}\Big]$$

Hence

$$\begin{gathered}\boldsymbol{E}\big[\,h(\boldsymbol{N})\,\big]-\boldsymbol{E}\big[\,h(\boldsymbol{F})\,\big]=\Psi(1)-\Psi(0)=\int_0^1\Psi'(t)dt\\=\frac{1}{2}\sum_{i,j=1}^{d}\int_0^1\mathbb{E}\Big[\,\partial^2_{x_ix_j}h\big(\,\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N}\,\big)\Big(\,C_{ij}-\big(\,DF_j,-DL^{-1}F_i\,\big)_{\mathfrak{X}}\,\Big)\,\Big].\end{gathered}\tag{3.4.14}$$

□

We now have (almost) all the information we need to prove the following remarkable result.

Theorem 3.4.12. *Let $d\geq 2$ and $q_1,\dots,q_d\in\mathbb{N}$. Consider the d-dimensional random vector*

$$\boldsymbol{F}=(F_1,\dots,F_d),\;\;F_i\in\mathfrak{X}^{:q_i:},\;\;i=1,\dots,d.$$

Let $f_i\in\mathfrak{X}^{\odot q_i}$ such that $\boldsymbol{I}_{q_i}[f_i]=F_i$. Denote by C the covariance matrix of the random vector $\boldsymbol{F}$, $C_{ij}=\mathbb{E}[F_iF_j]$, and let $\boldsymbol{N}\sim\mathcal{N}(0,C)$. Consider the continuous function

$$\psi:(\mathbb{R}\times\mathbb{R}_{>0})^d\to\mathbb{R}>0$$

given by

$$\Psi(x_1,y_1,\dots,x_d,y_d)=\sum_{i,j=1}^{d}\delta_{q_iq_j}\left(\sqrt{\sum_{r=1}^{q_i-1}\binom{2r}{r}}\right)|x_i|^{\frac{1}{2}}$$

$$+\sum_{i,j=1}^{d}(1-\delta_{q_iq_j})\left(\,\big(\,2|y_j|\,\big)^{\frac{1}{2}}\,|x_i|^{\frac{1}{4}}++\sum_{r=1}^{\min(q_i,q_j)-1}\sqrt{(2(q_i+q_j-2r)!}\binom{q_j}{r}|x_i|^{\frac{1}{2}}\right),$$

and set

$$m(\boldsymbol{F})=\psi\big(\,m_4(F_1)-3m_2(F_1)^2,m_2(F_1),\dots,m_4(F_d)-3m_2(F_d)^2,m_2(F_d)\,\big),$$

where we recall that $m_k(X)$ denotes the k-th moment of a random variable X. Note that

$$\psi\big(x_1, y_1, \dots, x_d, y_d\big)_{x_1=\cdots=x_d=0} = 0.$$

If $h : \mathbb{R}^d \to \mathbb{R}$ is a C^2 function with bounded second derivatives, then

$$\big|\,\mathbb{E}\big[\,h(\boldsymbol{F})\,\big] - \mathbb{E}\big[\,h(N)\,\big]\,\big| \leq \frac{1}{2}\|h''\|_\infty m(\boldsymbol{F}).$$

The main ideas. We plan to use Proposition 3.4.10 so we need to estimate from above the quantities

$$\mathbb{E}\Big[\,\Big(\,C_{ij} - \big(\,DF_j, -DL^{-1}F_i\big)_{\mathfrak{X}}\,\Big)^2\,\Big] = \mathbb{E}\Big[\,\Big(\,\mathbb{E}[F_iF_j] - \frac{1}{q_i}\big(\,DF_i, DF_j\big)_{\mathfrak{X}}\,\Big)^2\,\Big].$$

Note that $C_{ij} = 0$ if $q_i \neq q_j$. Thus, we need to produce suitable upper estimates for quantities of the form

$$\mathbb{E}\Big[\,\alpha - \frac{1}{p}(DF, DG)_{\mathfrak{X}}\,\Big],\;\; F \in \mathfrak{X}^{:p:},\;\; G \in \mathfrak{X}^{:q:},\;\; \alpha \in \mathbb{R}.$$

This is what the next lemma accomplishes.

Lemma 3.4.13. *Let $F = \mathfrak{I}_p[f]$, $f \in \mathfrak{X}^{\odot p}$ and $G = \mathfrak{I}_q[g]$, $g \in \mathfrak{X}^{\odot q}$, $p, q \geq 1$. Suppose that α is a real constant.*

(i) *If $p = q$, then*

$$\mathbb{E}\Big[\,\Big(\,\alpha - \frac{1}{p}(DF, DG)_{\mathfrak{X}}\,\Big)^2\,\Big] \leq \big(\,\alpha - \mathbb{E}[FG]\,\big)^2 + \frac{p^2}{2}\sum_{r=1}^{p-1}(r-1)!\binom{p-1}{r-1}^4(2p-2r)!\big(\,\|f \boxtimes_{p-r} f\|^2_{\mathfrak{X}^{\otimes 2r}} + \|g \boxtimes_{p-r} g\|^2_{\mathfrak{X}^{\otimes 2r}}\,\big) \tag{3.4.15}$$

(ii) *If $p < q$, then*

$$\mathbb{E}\Big[\,\Big(\,\frac{1}{q}(DF, DG)_{\mathfrak{X}}\,\Big)^2\,\Big] \leq (p!)^2\binom{q-1}{p-1}^2(q-p)!\|f\|^2_{\mathfrak{X}^{\otimes p}}\|\,g \boxtimes_{q-p} g\|_{\mathfrak{X}^{\otimes 2p}} + \frac{p^2}{2}\sum_{r=1}^{p-1}(r-1)!\binom{p-1}{r-1}^2\binom{q-1}{r-1}^2(p+q-2r)!\big(\,\|f \boxtimes_{p-r} f\|^2_{\mathfrak{X}^{\otimes 2r}} + \|g \boxtimes_{q-r} g\|^2_{\mathfrak{X}^{\otimes 2r}}\,\big). \tag{3.4.16}$$

Main idea of the proof. The lemma follows from the identity

$$\big(DF, DG\big)_{\mathfrak{X}} = pq\sum_{r=1}^{\min(p,q)}(r-1)!\binom{p-1}{r-1}\binom{q-1}{r-1}\mathfrak{I}_{p+q-2r}[f \boxtimes_r g],$$

which can be reduced to the equality (3.1.29). For details we refer to [**124**, Lemma 6.2.1]. □

Using (3.4.7c) we deduce that for any $q \geq 2$ and any $f \in \mathfrak{X}^{\odot q}$ we have

$$\|f \otimes_r f\|^2_{\mathfrak{X}^{\otimes(2q-2r)}} \leq \frac{\big(\,r!(q-r)!\,\big)^2}{(q!)^4}\Big(\,\mathbb{E}\big[\,\boldsymbol{I}_q[\,f\,]^4\,\big] - 3\mathbb{E}\big[\,\boldsymbol{I}_q[\,f\,]^2\,\big]^2\,\Big).$$

Theorem 3.4.12 now follows from the above lemma after some simple algebraic manipulations □

Theorem 3.4.12 implies the following remarkable result.

Theorem 3.4.14 (Peccati-Tudor, [**132**]). *Let $d \geq 1$ and $q_1, \dots, q_d \in \mathbb{N}$. Consider the sequence of d-dimensional random vectors*

$$\boldsymbol{F}_n = (F_{1,n}, \dots, F_{d,n}), \;\; F_{j,n} \in \mathfrak{X}^{:q_j:}, \;\; j = 1, \dots, d, \;\; n \in \mathbb{N}.$$

Suppose that $C \in \mathcal{L}(\mathbb{R}^d)$ is symmetric and nonnegative definite and

$$\lim_{n\to\infty} \mathbb{E}\big[\, F_{i,n} F_{j,n} \,\big] = C_{ij}, \;\; \forall i, j = 1, \dots, d.$$

Then the following statements are equivalent.

(i) *The random vector $\boldsymbol{F}_n$ converges in probability to a Gaussian vector $\boldsymbol{N} \sim \mathcal{N}(0, C)$.*

(ii) *For each $j = 1, \dots, d$ the sequence of random variables $\big(F_{i,n}\big)_{n\in\mathbb{N}}$ converges in probability to a Gaussian r.v. $N_i \sim \mathcal{N}(0, C_{ii})$.*

□

The above result leads to the following substantial strengthening of Proposition 3.4.1

Theorem 3.4.15. *Consider a sequence of random variables $(F_\nu)_{\nu\geq 1}$ in $L^2_{\mathfrak{X}}\big(\Omega\big)$ such that $\mathbb{E}[F_\nu] = 0$, $\forall \nu$, i.e., $\mathrm{Proj}_0(F_\nu) = 0$, $\forall \nu$. Suppose that the following hold.*

(C_1) *For any $k \in \mathbb{N}$, $\exists v_k \geq 0$ such that*

$$\lim_{\nu\to\infty} \mathbb{E}\Big[\, \Big(\mathrm{Proj}_k\, F_\nu)\Big)^2 \,\Big] = \delta_{jk} v_k.$$

(C_2) *The sequence*

$$V_N := \sup_{\nu\geq 1} \sum_{k>N} \mathbb{E}\Big[\, \Big(\mathrm{Proj}_k\, F_\nu\Big)^2 \,\Big]$$

converges to 0 as $N \to \infty$.

(C_3') *For any $k \in \mathbb{N}$*

$$\lim_{\nu\to\infty} \mathbb{E}\Big[\, \Big(\mathrm{Proj}_k\, F_\nu\Big)^4 \,\Big] = 3v_k^2.$$

(C_3'')

Then the following hold.

(i) *The series $\sum_{n\geq 1} v_\nu$ is convergent. We denote by v its sum.*

(ii)

$$\lim_{\nu\to\infty} \mathrm{Var}(F_\nu) = v.$$

(iii) *As $\nu \to \infty$, the random variable F_ν converges in law to a random variable $F_\infty \sim \mathcal{N}(0, v)$.*

Remark 3.4.16. (a) The fourth moment theorem (Corollary 3.4.8) shows that the conditions $C_1 + C_3'$ are equivalent with the requirement that, $\forall p \in \mathbb{N}$, as $\nu \to \infty$ the random variables

$$\mathrm{Proj}^{\leq p}(F_\nu) = \sum_{k=1}^{p} \mathrm{Proj}_k[F_\nu]$$

converge in probability as $\nu \to \infty$ to a normal random variable with mean zero and variance $v_1 + \cdots + v_p$. This is condition (C_3) in Proposition 3.4.1.

(b) If we write

$$\mathrm{Proj}_k[F_\nu] = \mathfrak{I}_k[f_{\nu,k}], \quad f_{\nu,k} \in \mathfrak{X}^{:k:},$$

the Corollary 3.4.8 shows that the condition C_3'' is equivalent to

$$\lim_{\nu\to\infty} \|f_{\nu,k} \tilde{\boxtimes}_r f_{\nu,k}\|_{\mathfrak{X}^{\odot(2q-2r)}} = 0, \quad \forall k \geq 1, \quad \forall r = 1\dots, k-1$$

□

3.5. The number of critical points of Gaussian functions on Euclidean spaces

Suppose that $\mathfrak{a} : \mathbb{R} \to \mathbb{R}$ is an even Schwartz function such that $\mathfrak{a}(0) = 1$. Consider the isotropic Gaussian function $\Phi_\mathfrak{a}$ defined in Example 1.2.35.

More precisely consider the finite Borel measure $\mu \in \mathrm{Meas}(\mathbb{R}^m)$

$$\mu\big[\, d\xi \,\big] = \mu_\mathfrak{a}\big[\, d\xi \,\big] = \frac{1}{(2\pi)^m} w_{\mathfrak{a},m}\big(\, \xi \,\big)\boldsymbol{\lambda}\big[\, d\xi \,\big], \quad w_{\mathfrak{a},m}\big(\, \xi \,\big) = \mathfrak{a}\big(\, |\xi| \,\big)^2.$$

Its characteristic function is the nonnegative definite function defined by (1.2.30),

$$\boldsymbol{K}\big(\, \boldsymbol{x} \,\big) = \boldsymbol{K}_\mathfrak{a}\big(\, \boldsymbol{x} \,\big) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \xi, \boldsymbol{x}\rangle} \mu_\mathfrak{a}\big[\, d\xi \,\big] = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \xi, \boldsymbol{x}\rangle} \mathfrak{a}\big(\, |\xi| \,\big)^2 \boldsymbol{\lambda}\big[\, d\xi \,\big].$$

The function $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{K}_\mathfrak{a}\big(\, \boldsymbol{x} - \boldsymbol{y} \,\big)$ is the covariance kernel of $\Phi_\mathfrak{a}$.

For $R > 0$ we set

$$\mathfrak{a}_R(t) := \mathfrak{a}\big(t/R\big), \quad \forall t \in \mathbb{R}.$$

Consider the finite Borel measure $\mu_\mathfrak{a}^R \in \mathrm{Meas}(\mathbb{R}^m)$

$$\mu_\mathfrak{a}^R\big[\, d\xi \,\big] = \frac{1}{(2\pi)^m} w_{\boldsymbol{a}_R,m}\big(\, \xi \,\big)\boldsymbol{\lambda}\big[\, d\xi \,\big] = \frac{1}{(2\pi)^m} \mathfrak{a}\big(\, |\xi|/R \,\big)^2 \boldsymbol{\lambda}\big[\, d\xi \,\big].$$

Note that as $R \nearrow \infty$ the measure $\mu_\mathfrak{a}^R$ converges vaguely to $(2\pi)^{-m}\boldsymbol{\lambda}$. Its characteristic function is the nonnegative definite function

$$\boldsymbol{K}_\mathfrak{a}^R\big(\, \boldsymbol{x} \,\big) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \xi, \boldsymbol{x}\rangle} \mathfrak{a}\big(\, |\xi|/R \,\big)^2 d\xi = R^m \boldsymbol{K}_\mathfrak{a}\big(\, R\boldsymbol{x} \,\big). \tag{3.5.1}$$

We deduce that $\boldsymbol{K}_\mathfrak{a}^R\big(\, \boldsymbol{x} - \boldsymbol{y} \,\big)$ is the covariance kernel of the Gaussian function

$$\Phi_\mathfrak{a}^R(\boldsymbol{x}) := R^{m/2} \Phi_\mathfrak{a}\big(\, R\boldsymbol{x} \,\big).$$

Chapter 4

Random Morse functions on compact Riemann manifolds

In this last and main chapter of the book I will discuss the main character that interested me all along. To any amplitude $\mathfrak{a}$ and Riemann metric g on a smooth compact manifold M we will canonically associate a smooth random function $F^g_{\mathfrak{a}}$.

Fix the metric g and look at its "inflations" $g(R) = R^2 g$, $R \nearrow \infty$. The metric $g(R)$ is flatter and flatter as $R \nearrow \infty$. and we set $F^R_{\mathfrak{a}} := F^{g(R)}_{\mathfrak{a}}$. When M is the m-th dimensional torus and g is the canonical flat metric of volume 1, these functions coincide with the random functions $F^R_{\mathfrak{a}}$ constructed in Example 1.2.31.

To $F^R_{\mathfrak{a}}$ we associate the random measure $\mathfrak{C}^R_{\mathfrak{a}}$ defined by

$$\mathfrak{C}^R_{\mathfrak{a}}[f] = \sum_{dF^R_{\mathfrak{a}}(x)=0} f(x),$$

for any bounded Borel measurable function $f : M \to \mathbb{R}$. One of the main goals of this chapter is to prove a universality theorem stating that as $\mathbb{N} \ni N \to \infty$ the random measures $\frac{1}{N^m}\mathfrak{C}^N_{\mathfrak{a}}$ converge a.s. and L^2 to the *deterministic measure* $C_m(\mathfrak{a})\,\mathrm{vol}_g$ where vol_g denotes the volume element on M determined by g. We view it as a finite Borel measure on M. The constant $C_m(\mathfrak{a})$ is described in (2.3.26). When $M = \mathbb{T}^m$ and g is the canonical flat metric of volume 1 this result specializes to our earlier Law of Large Numbers, Corollary 2.5.10.

The Gaussian function $F^R_{\mathfrak{a}}$ has the same critical set as the function $R^{m/2}F^R_{\mathfrak{a}}$. On the other hand the random function $R^{m/2}F^R_{\mathfrak{a}}$ converges to the Gaussian noise on M associated to vol_g and for this reason we will refer to the large R limit as white noise limit.

4.1. A family of random smooth functions on a compact Riemann manifold

4.1.1. By way of motivation. Suppose that (M, g) is a smooth, compact, connected m-dimensional manifold. The Riemann metric allows us to define various concepts of random functions. The volume elementvol$_g$ determines various L^p-spaces, $L^p\big(M, g\big) = L^p\big(M, \mathrm{vol}_g\big)$, $1 \leq p \leq \infty$.

The metric g defines a Laplacian $\Delta_g : C^\infty(M) \to C^\infty(M)$. It has an unbounded discrete spectrum Spec $\big(\Delta_g\big)$ consisting only of nonnegative eigenvalues with finite multiplicities

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots .$$

In the above sequence each eigenvalues appears as often as its multiplicity. We get a complete orthonormal family of eigenfunctions $(\psi_n)_{n\geq 0}$

$$\Delta_g \psi_n = \lambda_n \psi_n, \;\; \|\psi_n\|_{L^2(M,g)} = 1.$$

Any smooth function $f \in C^\infty(M)$ admits a Fourier decomposition

$$f = \sum_{n\geq 0} c_n \psi_n.$$

We would like to sample $C^\infty(M)$ without any bias. One reasonable way to do this would be to require that the Fourier coefficients c_n be i.i.d. random variables. Here we encounter the first problem: if the coefficients are i.i.d. random variables the resulting random series would not be convergent to a function.

There is another less difficult issue. The eigenvalues could have multiplicity. For any $\lambda \geq 0$ we denote by Proj_λ the orthogonal projection onto $H_\lambda := \ker(\Delta_g - \lambda)$. We can rewrite the Fourier decomposition of f in an invariant way,

$$f = \sum_{\lambda\in \mathrm{Spec}(\Delta_g)} f_\lambda, \;\; f_\lambda = \mathrm{Proj}_\lambda f.$$

If f is random, then each component f_λ will be a random element of the eigenspace H_λ. For this problem to be geometrically well posed, the distribution of f_λ should be independent of the orthonormal basis of H_λ, i.e., it has to be invariant with respect to the orthogonal transformations of this finite dimensional Euclidean space. We will assume that the distribution of f_λ is Gaussian and its variance has the form $w(\lambda)\mathbb{1}_{H_\lambda}$.

If we choose $w(\lambda) = 1$, $\forall \lambda \in$ Spec $\big(\Delta_g\big)$, we again face the issue of nonconvergent random series. We can try the next best thing and set $w(\lambda) = 1$ for $\lambda \leq R^2$, where $R \gg 0$. We obtain a random function

$$F^R = \sum_{\lambda_n \leq R^2} X_n \psi_n,$$

$(X_n)_{n\geq 0}$ is a sequence of independent standard normal random variables. Then

$$F^R \in H_{\leq R^2} = \bigoplus_{\lambda \leq R^2} H_\lambda$$

The distribution of F^R is the canonical Gaussian measure on the Euclidean space $H_{\leq R}$. Note that the critical set of F^R is coincides with that of the normalized random function

$$\overline{F}^R := \frac{1}{\|F^R\|_{L^2(M,g)}} F^R,$$

and the random vector $\overline{F}^R$ is uniformly distributed on the unit sphere of $H_{\leq R^2}$.

Hence, as far as the distribution of critical points is concerned, the random function F^R samples the space $H_{\leq R^2}$ without any bias. As $R \to \infty$, the random function F^R samples unbiasedly more and more of the smooth function and in the limit we get a nondiscriminatory taste of all the Morse functions on M.

Weyl's asymptotic formula with Hörmander's error estimate [**73**] shows that

$$\dim H_{\leq R^2} = C_m(g)R^m + O\big(\, R^{m-1}\,\big) \;\text{ as } R \to \infty, \tag{4.1.1}$$

where

$$C_m(g) := \frac{\boldsymbol{\omega}_m}{(2\pi)^m}\,\mathrm{vol}_g(M).$$

This estimate implies that (see [**145**, Prop. 13.1])

$$\lambda_n \sim C_m(g)^{-\frac{2}{m}} n^{\frac{2}{m}} \;\text{ as } n \to \infty. \tag{4.1.2}$$

Appendix A

Differential geometry

A.1. Jacobians and the coarea formula

.

At its core, the coarea formula is a sophisticated version of Fubini's Theorem. To best understand this we begin with the simplest case.

Recall Fubini's theorem. Suppose φ is a integrable function on $\mathbb{R}^{n+k}$. Then

$$\begin{aligned}&\int_{\mathbb{R}^{n+k}} \varphi(x^1,\dots,x^{n+k})dx^1\cdots dx^{n+k}\\ &=\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^k}\varphi(x^1,\dots,x^n,x^{n+1},\dots,x^{n+k})dx^{n+1}\cdots dx^{n+k}\right)dx^1\cdots dx^n.\end{aligned}$$

We can reformulate this as follows. Set

$$\boldsymbol{y}=(x^1,\dots,x^n),\;\;\boldsymbol{x}=(x^{n+1},\dots,x^{n+k})$$

and define $A:\mathbb{R}^{n+k}\to\mathbb{R}^n,\;\;(\boldsymbol{x},\boldsymbol{y})\mapsto\boldsymbol{y}$. Then

$$\int_{\mathbb{R}^{n+k}}\varphi(\boldsymbol{x},\boldsymbol{y})\operatorname{vol}_{n+k}[d\boldsymbol{x}d\boldsymbol{y}]=\int_{\mathbb{R}^n}\left(\int_{A^{-1}(\boldsymbol{y})}\varphi(\boldsymbol{x},\boldsymbol{y})\operatorname{vol}_k[d\boldsymbol{x}]\right)\operatorname{vol}_n[d\boldsymbol{y}]. \tag{A.1.1}$$

where vol_i denotes the i-dimensional Lebesgue measure.

Consider now a slightly more general case of a linear map

$$A:\mathbb{R}^{n+k}\to\mathbb{R}^n,\;\;(x^1,\dots,x^n,x^{n+1},\dots,x^{n+k})\mapsto(y^1,\dots,y^n)=(\mu_1x^1,\dots,\mu_nx^n), \tag{A.1.2}$$

where $\mu_1,\dots,\mu_n$ are positive numbers. Applying the Fubini theorem we deduce

$$\begin{aligned}&\int_{\mathbb{R}^{n+k}}\mu_1\cdots\mu_n\varphi(x^1,\dots,x^{n+k})\operatorname{vol}_{n+k}[dx^1\cdots dx^{n+k}]\\ &=\int_{\mathbb{R}^{n+k}}\varphi\Big(\frac{y^1}{\mu^1},\dots\frac{y^n}{\mu^n},x^{n+1},\dots,x^{n+k}\Big)\operatorname{vol}_{n+k}[dy^1\cdots dy^ndx^{n+1}\cdots dx^{n+k}]\\ &=\int_{\mathbb{R}^n}\left(\int_{A^{-1}(\boldsymbol{y})}\varphi(\boldsymbol{x},\boldsymbol{y})\operatorname{vol}_k[d\boldsymbol{x}]\right)\operatorname{vol}_n[d\boldsymbol{y}].\end{aligned} \tag{A.1.3}$$

But for the factor $\mu_1 \cdots \mu_n$, the formulæ (A.1.1) and (A.1.3) look similar. To give an invariant meaning to this quantity we need to use the following elementary fact of linear algebra.

Lemma A.1.1. *Suppose that* $\boldsymbol{U}$ *and* $\boldsymbol{V}$ *are* Euclidean *spaces, respectively of dimensions* $n+k$ *and* n $(n, k \geq 0)$, *and* $A : \boldsymbol{U} \to \boldsymbol{V}$ *is a linear map. Then there exist Euclidean coordinates* $x^1, \dots, x^{n+k}$ *on* $\boldsymbol{U}$, *Euclidean coordinates* $y^1, \dots, y^n$ *on* $\boldsymbol{V}$ *and nonnegative numbers* $\mu_1, \dots, \mu_n$ *such that, in these coordinates the operator* A *is described by*

$$y^j = \mu_j x^j, \;\; 1 \leq j \leq n.$$

The numbers $\mu_1^2, \dots, \mu_n^2$ *are the eigenvalues of the positive symmetric operator* $AA^* : \boldsymbol{V} \to \boldsymbol{V}$ *so that*

$$\mu_1 \cdots \mu_n = J_A := \sqrt{\det AA^*}.$$

In particular

$$A \;\; \textit{surjective} \Longleftrightarrow J_A \neq 0.$$

The quantity J_A *is called the* Jacobian *of the linear map* A. □

Thus, we can rewrite (A.1.3) as

$$\int_{\mathbb{R}^{n+k}} J_A(\boldsymbol{x}, \boldsymbol{y}) \varphi(\boldsymbol{x}, \boldsymbol{y}) \operatorname{vol}_{n+k}[d\boldsymbol{x} d\boldsymbol{y}] = \int_{\mathbb{R}^n} \left(\int_{A^{-1}(\boldsymbol{y})} \varphi(\boldsymbol{x}, \boldsymbol{y}) \operatorname{vol}_{A^{-1}(\boldsymbol{y})}[d\boldsymbol{x}] \right) |dV_n(\boldsymbol{y})|, \tag{A.1.4}$$

where $\operatorname{vol}_{A^{-1}(\boldsymbol{y})}$ denotes the Eucldean volme element on the affine subspace $A^{-1}(\boldsymbol{y})$. Lemma A.1.1 shows that (A.1.4) holds for any *surjective linear* map $\mathbb{R}^{n+k} \to \mathbb{R}^n$.

Proposition A.1.2. *Suppose that* $\boldsymbol{U}$ *and* $\boldsymbol{V}$ *are* Euclidean *spaces, respectively of dimensions* $n+k$ *and* n $(n, k \geq 0)$, *and* $A : \boldsymbol{U} \to \boldsymbol{V}$ *is a linear map. Then*

$$J_A = \frac{\operatorname{vol}_n \left[A(B_1^{\boldsymbol{U}}) \right]}{\operatorname{vol}_n \left[B_1^{\boldsymbol{V}} \right]}, \tag{A.1.5}$$

where $B_1^{\boldsymbol{U}}$ *denotes the unit ball in* $\boldsymbol{U}$ *and* $B_1^{\boldsymbol{V}}$ *the unit ball in* $\boldsymbol{V}$.

Proof. Choose coordinates (x^i) on $\boldsymbol{U}$ and (y^j) on $\boldsymbol{V}$ as in Lemma A.1.1. If A is not onto the result is obvious since, then $\dim A(\boldsymbol{U}) < n$. If A is onto, then $A(B_1^{\boldsymbol{U}})$ is isometric to the ellipsoid

$$E = \left\{ x \in \mathbb{R}^n; \;\; \sum_{j=1}^n \frac{(x^j)^2}{\mu_j^2} \leq 1 \right\}$$

where the numbers μ_j are as in Lemma A.1.1. Observe that $\operatorname{vol}_n \left[E \right] = \mu_1 \cdots \mu_n$. □

Remark A.1.3. Suppose that $k = 0$ so $\dim \boldsymbol{U} = \dim \boldsymbol{V} = n$. Assume that A is onto. Then the push-forward by A of the Lebesgue measure on $\boldsymbol{U}$ is given by

$$A_{\#} \boldsymbol{\lambda}_{\boldsymbol{U}} = \frac{1}{J_A} \boldsymbol{\lambda}_{\boldsymbol{V}}. \tag{A.1.6}$$

If $\boldsymbol{U}$ and $\boldsymbol{V}$ are equipped with orientations, then we can invariantly define $\det A$ and we have $J_A = |\det A|$. □

It is convenient to give a more explicit algebraic description of J_A. This relies on the concept of Gram determinant. More precisely, given a collection of vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ in an Euclidean space $\boldsymbol{U}$ we define their *Gram determinant* (or *Gramian*) to be the quantity

$$\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) := \det\Big((\boldsymbol{u}_i, \boldsymbol{u}_j)_{\boldsymbol{U}} \Big)_{1\leq i,j\leq n},$$

where $(-,-)_{\boldsymbol{U}}$ denotes the inner product in $\boldsymbol{U}$. Geometrically, $\sqrt{\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)}$ is the n-dimensional volume of the parallelepiped spanned by the vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$,

$$P(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n) = \Big\{ \sum_{j=1}^n t_j \boldsymbol{w}_j; \;\; t_j \in [0,1] \Big\}.$$

Note that $\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = 0$ iff the vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ are linearly dependent and $\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = 1$ if the vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ form an orthonormal system.

Equivalently

$$\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = \big(\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n, \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_n \big)_{\Lambda^n \boldsymbol{U}}$$

where $(-,-)_{\Lambda^n U}$ denotes the inner product on $\Lambda^n \boldsymbol{U}$ induced by the inner product in $\boldsymbol{U}$.

Lemma A.1.4. *Let $A : \boldsymbol{U} \to \boldsymbol{V}$ be as in Lemma A.1.1. Fix a basis $\boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k}$ of $\boldsymbol{U}_0 := \ker A$ and vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ such that $A\boldsymbol{u}_1, \dots, A\boldsymbol{u}_n$ span $\boldsymbol{V}$. Then*

$$J_A^2 = \frac{\mathbb{G}(A\boldsymbol{u}_1, \dots A\boldsymbol{u}_n)\mathbb{G}(\boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k})}{\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k})}. \tag{A.1.7}$$

Proof. We first prove the result when $\dim \boldsymbol{U} = \dim \boldsymbol{V}$. In this case the collection $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ is a basis of $\boldsymbol{U}$. Fix an orthonormal basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ of $\boldsymbol{U}$ denote by $T : \boldsymbol{U} \to \boldsymbol{U}$ the linear operator $\boldsymbol{e}_j \mapsto \boldsymbol{u}_j$. Then

$$\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = \det T^*T,$$

$$\mathbb{G}(A\boldsymbol{u}_1, \dots A\boldsymbol{u}_n) = \det((AT)^*(AT)) = |\det T^*| \det AA^* |\det T| = J_A^2 \det TT^*.$$

To deal with the general case, we denote by P_0 the orthogonal projection onto $\boldsymbol{U}_0$. Now define

$$\widehat{A} : \boldsymbol{U} \to \widehat{\boldsymbol{V}} := \boldsymbol{V} \oplus \boldsymbol{U}_0, \;\; \boldsymbol{u} \mapsto A\boldsymbol{u} \oplus P_0\boldsymbol{u}.$$

we equip $\widehat{\boldsymbol{V}}$ with the product Euclidean structure.

Let us observe that $J_A = J_{\widehat{A}}$. Indeed, with respect to the (orthogonal) direct sum decomposition $\widehat{\boldsymbol{V}} = \boldsymbol{V} \oplus \boldsymbol{U}_0$ the operator $\widehat{A}\widehat{A}^*$ has the block decomposition

$$\widehat{A}\widehat{A}^* = \left[\begin{array}{cc} AA^* & 0 \\ * & \mathbb{1}_{\boldsymbol{U}_0} \end{array} \right]$$

so that

$$\det \widehat{A}\widehat{A}^* = \det AA^*.$$

Observe that in $\Lambda^{n+k}(\boldsymbol{V} \oplus \boldsymbol{U}_0)$ we have the equality

$$\hat{A}\boldsymbol{u}_1 \wedge \cdots \hat{A}\boldsymbol{u}_n \wedge \boldsymbol{f}_{n+1} \wedge \cdots \wedge \boldsymbol{f}_{n+k} = A\boldsymbol{u}_1 \wedge \cdots A\boldsymbol{u}_n \wedge \boldsymbol{f}_{n+1} \wedge \cdots \wedge \boldsymbol{f}_{n+k}$$

so that

$$\begin{aligned} \mathbb{G}(\widehat{A}\boldsymbol{u}_1, \dots, \widehat{A}\boldsymbol{u}_n, \widehat{A}\boldsymbol{f}_{n+1}, \dots, \widehat{A}\boldsymbol{f}_{n+k}) &= \mathbb{G}(A\boldsymbol{u}_1, \dots, A\boldsymbol{u}_n, \boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k}) \\ &= \mathbb{G}(A\boldsymbol{u}_1, \dots A\boldsymbol{u}_n)\mathbb{G}(\boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k}). \end{aligned}$$

Now apply the first part of the proof to deduce that

$$J_A^2 = J_{\widehat{A}}^2 = \frac{\mathbb{G}(\widehat{A}\boldsymbol{u}_1, \dots, \widehat{A}\boldsymbol{u}_n, \widehat{A}\boldsymbol{f}_{n+1}, \dots, \widehat{A}\boldsymbol{f}_{n+k})}{\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k})} = \frac{\mathbb{G}(A\boldsymbol{u}_1, \dots A\boldsymbol{u}_n)\mathbb{G}(\boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k})}{\mathbb{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{f}_{n+1}, \dots, \boldsymbol{f}_{n+k})}.$$

□

Suppose now that X and Y are C^1 manifolds of dimensions $n+k$ and respectively n, $k \geq 0$ equipped with Riemann metrics g_X and g_Y. We denote by vol_X and vol_Y the volume measures induced by g_X and respectively g_Y. Let $\Phi : X \to Y$ is a C^1-map. For $x \in X$ we denote by $\Phi'(x)$ the differential of Φ at x. This is a linear map

$$\Phi'(u) : T_x x \to T_{F(u)} Y.$$

The *Jacobian* of the map Φ is the function

$$J_\Phi : X \to [0, \infty), \;\; J_\Phi(x) = \sqrt{\det\big(\,\Phi'(x)\Phi'(x)^*\,\big)},$$

where $\Phi'(x)^* : T_{F(u)}Y \to T_u U$ is the adjoint of $\Phi'(x)$ determined by the inner products g_x^X on $T_x Y$ and $g_{\Phi(x)}^Y$ on $T_{\Phi(x)} Y$.

Theorem A.1.5 (The coarea formula: version 1)**.** *Suppose that $\Phi : X \to Y$ is a C^1-map such that for any $x \in X$ the differential $\Phi'(x)$ is surjective. We denote by $J_\Phi(x)$ the Jacobian of this map. For any nonnegative Borel measurable function $\varphi : X \to \mathbb{R}$ we have*

$$\int_X J_\Phi(x)\varphi(x)\,\mathrm{vol}_X\big[\,dx\,\big] = \int_Y \left(\int_{\Phi^{-1}(y)} \varphi(x)\,\mathrm{vol}_{\Phi^{-1}(y)}\big[\,dx\,\big] \right) \mathrm{vol}_Y\big[\,dy\,\big], \tag{A.1.8}$$

where $\mathrm{vol}_{\Phi^{-1}(y)}$ denotes the volume density on the fiber $\Phi^{-1}(y)$ induced by the restriction of g_X to $\Phi^{-1}(y)$.

Proof. We consider first the case when X is an open subset of $\mathbb{R}^{n+k}$ with coordinates $(x^1, \dots, x^{n+k})$ equipped with a C^1-metric g_X, Y is an open subset of $\mathbb{R}^k$ with coordinates $(y^1, \dots, y^k)$ equipped with a metric g_Y and the map Φ is given by

$$y^j = x^j, \;\; j = 1, \dots, n.$$

We have

$$\begin{aligned} \mathrm{vol}_X\big[\,dx\,\big] &= \sqrt{\mathbb{G}_X(\partial_{x^1}, \dots, \partial_{x^{n+k}})}\,\mathrm{vol}_{n+k}\big[\,dx^1 \cdots dx^{n+k}\,\big] \\ &= \underbrace{\sqrt{\mathbb{G}_X(\partial_{x^1}, \dots, \partial_{x^{n+k}})}}_{=:\rho_X}\,\mathrm{vol}_{n+k}\big[\,dy^1 \cdots dy^k dx^{k+1} \cdots dx^{n+k}\,\big], \end{aligned}$$

$$\mathrm{vol}_{\Phi^{-1}(y)}\big[\,dx^{n+1} \cdots dx^{n+k}\,\big] = \underbrace{\sqrt{\mathbb{G}_X(\partial_{x^{k+1}}, \dots, \partial_{x^{n+k}})}}_{=:\rho_\Phi}\,\mathrm{vol}_k\big[\,dx^{n+1} \cdots dx^{n+k}\,\big],$$

where the subscript X indicates that the inner product in the definition of the above Gramm determinants is the one determined by the Riemann metric on X. Similarly

$$\mathrm{vol}_Y\big[\,dy\,\big] = \underbrace{\sqrt{\mathbb{G}_Y(\partial_{y^1}, \dots, \partial_{y^n}))}}_{=:\rho_Y}\,\mathrm{vol}_n\big[\,dy\,\big] = \sqrt{\mathbb{G}_Y(\Phi'(x)\partial_{x^1}, \dots, \Phi'(x)\partial_{x^n}))}\,\mathrm{vol}_n\big[\,dy\,\big].$$

Using the Fubini theorem we deduce that for any nonnegative, measurable function $\phi : X \to \mathbb{R}$ we have

$$\int_X \rho_Y \phi \rho_X \operatorname{vol}_{n+k}\big[\, dy^1 \cdots dy^k dx^{k+1} \cdots dx^{n+k} \,\big]$$

$$= \int_Y \left(\int_{\Phi^{-1}(\boldsymbol{y})} \rho_X \phi \operatorname{vol}_k \big[\, dx^{n+1} \cdots dx^{n+k} \,\big] \right) \rho_Y \operatorname{vol}_n \big[\, dy^1 \cdots dy^n \,\big]$$

$$= \int_Y \left(\int_{\Phi^{-1}(y)} \frac{\rho_X}{\rho_\Phi} \phi \rho_\Phi \operatorname{vol}_k \big[\, dx^{n+1} \cdots dx^{n+k} \,\big] \right) \operatorname{vol}_Y \big[\, dy \,\big]$$

$$= \int_Y \left(\int_{\Phi^{-1}(\boldsymbol{y})} \frac{\rho_X}{\rho_\Phi} \phi \operatorname{vol}_{\Phi^{-1}(y)} \big[\, dx \,\big] \right) \operatorname{vol}_Y \big[\, dy \,\big].$$

Set $\varphi := \frac{\rho_Y \phi}{J_\Phi}$ so that $\phi = \frac{J_\Phi}{\rho_Y}\varphi$. Then the above equality can be rewritten

$$\int_X J_\Phi(x)\varphi(x) \operatorname{vol}_X \big[\, dx \,\big] = \int_Y \left(\int_{\Phi^{-1}(y)} \frac{\rho_X J_\Phi}{\rho_\Phi \rho_Y} \varphi(x) \operatorname{vol}_{\Phi^{-1}(y)} \big[\, dx \,\big] \right) \operatorname{vol}_Y \big[\, dy \,\big].$$

The co-area formula is proved once we show that

$$\frac{\rho_X J_\Phi}{\rho_\Phi \rho_Y} = 1, \;\; \text{i.e.,} \;\; J_\Phi = \frac{\rho_Y \rho_\Phi}{\rho_X}.$$

The last equality follows from (A.1.7).

The general case of the co-area formula can be reduced to the special case via partition of unity and the implicit function theorem. □

Corollary A.1.6. *Suppose X is a C^1 manifold equipped with a C^1-metric g_X, and $f : X \to \mathbb{R}$ is a C^1 function with no critical points. Then for any measurable function $\phi : X \to \mathbb{R}$ we have*

$$\int_X \phi(\boldsymbol{p}) |dV_X(\boldsymbol{p})| = \int_{\mathbb{R}} \left(\int_{\{f=t\}} \frac{\phi(\boldsymbol{p})}{|\nabla f(\boldsymbol{p})|} |dV_{f^{-1}(t)}(\boldsymbol{p})| \right) dt. \tag{A.1.9}$$

In particular, by setting $\phi = 1$ we deduce

$$\operatorname{vol}(X) = \int_{\mathbb{R}} \left(\int_{\{f=t\}} \frac{1}{|\nabla f(\boldsymbol{p})|} |dV_{f^{-1}(t)}(\boldsymbol{p})| \right) dt. \tag{A.1.10}$$

□

Example A.1.7. Consider the unit sphere

$$S^n = \Big\{ (x_0, x_1, \dots, x_n) \in \mathbb{R}^n; \;\; \sum_{j=0}^n x_j^2 = 1 \Big\}.$$

and a continuous function $u : \mathbb{R} \to \mathbb{C}$ Let S^n_* denote the unit sphere with the poles $x_0 = \pm 1$ removed. Denote by $\boldsymbol{\sigma}_m$ the volume of the unit m-dimensional sphere.

Consider $f : S^n_* \to \mathbb{R}$, $f(x_0, \dots, x_n) = x_0$. This function has no critical points on S^n_*. Let $\boldsymbol{p} \in S^n_*$ such that $f(\boldsymbol{p}) = x_0(\boldsymbol{p}) = t$. Denote by φ the angle between the radius $O\boldsymbol{p}$ and the x_0-axis. Note that

$$\cos\varphi = x_0 = t.$$

The gradient of f is the projection of ∂_{x_0} on the tangent plane $T_{\boldsymbol{p}}S^n$. We deduce that

$$|\nabla f(\boldsymbol{p})| = |\partial_{x_0}|\sin\varphi = (1-t^2)^{1/2}.$$

The level set $\{f=t\}$ is an $(n-1)$-dimensional sphere of radius $(1-t^2)^{1/2}$ and we deduce

$$\int_{\{f=t\}} \frac{u(t)}{|\nabla f(\boldsymbol{p})|}|dV_{f^{-1}(t)}(\boldsymbol{p})| = (1-t^2)^{-1/2}\mathrm{vol}\,(f=t) = \boldsymbol{\sigma}_{n-1}u(t)(1-t^2)^{\frac{n-2}{2}}.$$

Hence

$$\int_{S^n} u(x^0)dV_{S^n} = \boldsymbol{\sigma}_{n-1}\int_{-1}^1 u(t)(1-t^2)^{\frac{n-2}{2}}dt. \tag{A.1.11}$$

When $u=1$ we deduce

$$\boldsymbol{\sigma}_n = \boldsymbol{\sigma}_{n-1}\int_{-1}^1 (1-t^2)^{\frac{n-2}{2}}dt = 2\boldsymbol{\sigma}_{n-1}\int_0^1(1-t^2)^{\frac{n-2}{2}}dt$$

$(t=\sqrt{s})$

$$= \boldsymbol{\sigma}_{n-1}\int_0^1(1-s)^{\frac{n}{2}-1}s^{\frac{1}{2}-1}ds.$$

We have a classical identity

$$B(p,q) := \int_0^1 (1-x)^{p-1}x^{q-1}dx \stackrel{(B.1.4)}{=} \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},\ \ p.q>0.$$

Hence

$$\frac{\boldsymbol{\sigma}_n}{\boldsymbol{\sigma}_{n-1}} = \frac{\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{n+1}{2})}.$$

Using the equalities $\boldsymbol{\sigma}_0 = 2$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ we deduce

$$\boldsymbol{\sigma}_n = \frac{2\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})}.$$

We can obtain easily $\boldsymbol{\omega}_n$, the volume of the unit n-dimensional ball,

$$\boldsymbol{\omega}_n = \frac{1}{n}\boldsymbol{\sigma}_{n-1} = \frac{\pi^{\frac{n}{2}}}{\frac{n}{2}\Gamma(\frac{n}{2})} = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}. \tag{A.1.12}$$

□

Theorem A.1.5 can be substantially generalized. For a proof of the next result we refer to [**31**, Sec.3.3] or [**109**, Sec. 3]. We denote by $\mathcal{H}^M_d$ the d-dimensional Hausdorff measure on a Riemann manifold (M,g). If $m=\dim M$, then $\mathcal{H}^M_m = \mathrm{vol}_M$.

Theorem A.1.8 (The coarea formula: version 2)**.** *Suppose X and Y are connected, Riemann C^1-manifolds of dimensions $n+k$ and respectively n, where $k\geq 0$. If $\Phi: X\to Y$ is a C^1-map satisfying the Lipschitz condition*

$$d_Y\big(\Phi(x_1),\Phi(x_2)\big) \leq Ld_X(x_1,x_2),\ \ \forall x_1,x_2\in X,$$

Then for any nonnegative Borel measurable functions $\alpha: X\to\mathbb{R}$ and $\beta: Y\to\mathbb{R}$ such that α has compact support we have

$$\int_X J_\Phi(x)\alpha(x)\Phi^*\beta(x)\,\mathrm{vol}_X\big[\,dx\,\big] = \int_Y\left(\int_{\Phi^{-1}(y)}\alpha(x)\mathcal{H}^X_k\big[\,dx\,\big]\right)\beta(y)\,\mathrm{vol}_Y\big[\,dy\,\big]. \tag{A.1.13}$$

The two sides of the above equality are simultaneously finite or infinite. If $\dim X = \dim Y = n$, *then the above equality reads*

$$\int_X J_\Phi(x)\alpha(x)\Phi^*\beta(x)\,\mathrm{vol}_X\big[\,dx\,\big] = \int_Y \left(\sum_{\Phi(x)=y} \alpha(x)\right)\beta(y)\,\mathrm{vol}_Y\big[\,dy\,\big] \tag{A.1.14}$$

□

Analysis

B.1. The Gamma function

Definition B.1.1 (Gamma and Beta functions)**.** The *Gamma function* is the function

$$\Gamma : (0,\infty) \to \mathbb{R},\ \ \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt. \tag{B.1.1}$$

The *Beta function* is the function of two positive variables

$$B(x,y) := \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},\ \ x,y > 0. \tag{B.1.2}$$

□

We gather here a few basic facts about the Gamma and Beta functions used in the text. For proofs we refer to [**86**, Chap. 1] or [**159**, Chap. 12].

Proposition B.1.2. *The following hold.*

(i) $\Gamma(1) = 1$.

(ii) $\Gamma(x+1) = x\Gamma(x)$, $\forall x > 0$.

(iii) *For any* $n = 1, 2, \dots$ *we have*

$$\Gamma(n) = (n-1)!. \tag{B.1.3}$$

(iv) $\Gamma(1/2) = \sqrt{\pi}$.

(v) *For any* $x, y > 0$ *we have* Euler's formula

$$B(x,y) = \int_0^1 s^{x-1}(1-s)^{y-1} ds = \int_0^\infty \frac{u^{x-1}}{(1+u)^{x+y}} du. \tag{B.1.4}$$

(vi) *For any* $x \in (0,1)$ *we have*

$$B(x, 1-x) = \Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x} \tag{B.1.5}$$

□

The equality (iv) above reads

$$\sqrt{\pi} = \Gamma(1/2) = \int_0^\infty e^{-t} t^{-1/2} dt$$

$(t = x^2,\ t^{-1/2} = x^{-1}\ dt = 2xdx)$

$$= 2\int_0^\infty e^{-x^2} dx = \int_{-\infty}^0 e^{-x^2} dx + \int_0^\infty e^{-x^2} dx = \int_{-\infty}^\infty e^{-x^2} dx.$$

If we make the change in variables $x = \frac{s}{\sqrt{2}}$ so that $x^2 = \frac{s^2}{2}$ and $dx = \frac{1}{\sqrt{2}} ds$, then we deduce

$$\sqrt{\pi} = \frac{1}{\sqrt{2}} \int_{-\infty}^\infty e^{-\frac{x^2}{2}} dx.$$

From this we obtain the fundamental equality

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{x^2}{2}} dx = 1. \tag{B.1.6}$$

The function $\Gamma(x)$ grows very fast as $x \to \infty$. Its asymptotics is governed by the *Stirling's formula*

$$\Gamma(x+1) = x\Gamma(x) \sim \sqrt{2\pi x}\left(\frac{x}{e}\right)^x \ \text{ as } x \to \infty. \tag{B.1.7}$$

Note that for $n \in \mathbb{N}$ the above estimate reads

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \ \text{ as } n \to \infty. \tag{B.1.8}$$

There are very sharp estimates for the ratio

$$q_n = \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}.$$

More precisely we have (see [**58**, II.9])

$$\frac{1}{12n+1} < \log q_n < \frac{1}{12n}. \tag{B.1.9}$$

In other words

$$\log n! = n\log n - n + \frac{1}{2}\log n + \frac{1}{2}\log(2\pi) + O\big(n^{-1}\big), \ \text{ as } n \to \infty.$$

We denote by $\boldsymbol{\omega}_n$ the volume of the n-dimensional Euclidean unit ball

$$B^n := \big\{\, \boldsymbol{x} \in \mathbb{R}^n;\ \ \|\boldsymbol{x}\| \leq 1 \,\big\},\ \ \|\boldsymbol{x}\| = \sqrt{x_1^2 + \cdots + x_n^2},$$

and by $\boldsymbol{\sigma}_{n-1}$ the "area" of the unit sphere in $\mathbb{R}^n$

$$S^{n-1} = \big\{\, \boldsymbol{x} \in \mathbb{R}^n;\ \ \|\boldsymbol{x}\| = 1 \,\big\}.$$

Then

$$\boldsymbol{\sigma}_{n-1} = \frac{2\Gamma(1/2)^n}{\Gamma(n/2)},\ \ \boldsymbol{\omega}_n = \frac{1}{n}\boldsymbol{\sigma}_{n-1} = \frac{\Gamma(1/2)^n}{\Gamma\big((n+1)/2\big)}. \tag{B.1.10}$$

B.2. The Fourier transform and tempered distributions.

Although the Fourier transform is a well known concept, there is quite a bit of variability in the conventions of various authors and we have included this section to make sure the reader is aware of the conventions we use. For proofs and more details we refer to [**50**, Chap.14] that served as our main source.

For $\alpha \in \big(\mathbb{Z}_{\geq 0}\big)^m$ we set

$$|\alpha| = \alpha_1 + \cdots + \alpha_m, \;\; \boldsymbol{x}^\alpha = x_1^{\alpha_1} \cdots x_m^{\alpha_m}, \;\; \partial^\alpha = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_m}^{\alpha_m}.$$

For any smooth complex valued function u and $k, N \in \mathbb{N}$ we set

$$p_{k,N}(f) = \sup_{\boldsymbol{x}\in\mathbb{R}^n} \big(1 + |x|^k\big) \sum_{|\alpha|\leq N} \big|\,\partial^\alpha u(\boldsymbol{x})\,\big|.$$

We denote by $\mathcal{S}(\mathbb{R}^m)$ the space of *Schwartz functions*, i.e., *complex valued* functions $u \in C^\infty(\mathbb{R}^m)$ such that $p_{k,N}(u) < \infty$, $\forall k, N \in \mathbb{N}$.

The countable collection of seminorms $\big(p_{k,N}\big)_{k,N\in\mathbb{N}}$ equips $\mathcal{S}(\mathbb{R}^m)$ with a structure of Fréchet space.

The Fourier transform is the linear map

$$\mathscr{F} : \mathcal{S}(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m), \;\; \mathscr{F}\big[\,u\,\big](\boldsymbol{\xi}) = \int_{\mathbb{R}^m} e^{-\boldsymbol{i}\langle \boldsymbol{\xi},\boldsymbol{x}\rangle} u(\boldsymbol{x}) d\boldsymbol{x}, \;\; \langle \boldsymbol{\xi}, \boldsymbol{x}\rangle := \sum_{j=1}^m \xi_j x_j.$$

We will frequently use the alternate notation $\widehat{u}($
$boxi) := \mathscr{F}\big[\,u\,\big](\boldsymbol{\xi})$. One can show that $\mathscr{F}$ is continuous with respect to the above Fréchet structure.

For $j = 1, \dots, m$ define $M_{x_j} : \mathcal{S}(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m)$, $M_{x_j}\big[\,u\,\big](x) = x_j u(x)$. Then

$$\mathscr{F} \circ M_{x_j} = \boldsymbol{i}\partial_{\xi_j} \circ \mathcal{F}, \;\; \mathcal{F} \circ \partial_{x_j} = \boldsymbol{i} M_{\xi_j} \circ \mathscr{F}. \tag{B.2.1}$$

Let $R : \mathcal{S}\big(\,(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m)$ denote the involution $Ru(x) = u(-x)$. The equalities (B.2.1) show that the operator $R \circ \mathscr{F} \circ \mathscr{F}$ commutes with M_{x_j} and ∂_{x_k} for any j, k. This can be used to show that there exists a constant c such that

$$R \circ \mathscr{F} \circ \mathscr{F} = c\mathbb{1}.$$

The equality (1.1.3) implies that $c = (2\pi)^{-m}$. Hence $\mathscr{F}$ is a bijection and its inverse is give by the *Fourier inversion formula*, $\mathscr{F}^{-1} = (2\pi)^{-m} R \circ \mathscr{F}$, i.e.,

$$u(x) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \xi, x\rangle} \widehat{u}(\xi) d\xi. \tag{B.2.2}$$

For $u, v \in \mathcal{S}(\mathbb{R}^m)$ we set

$$\langle u, v\rangle = \int_{\mathbb{R}^m} u(x)v(x)dx, \;\; \big(\,u, v\,\big) = \langle u, \bar{v}\rangle = \int_{\mathbb{R}^m} u(x)v(\bar{x})dx.$$

We then have the following fundamental equalities

$$\big\langle \widehat{u}, v\big\rangle = \big\langle u, \widehat{v}\big\rangle \tag{B.2.3a}$$

$$\big(\,u, v\,\big) = (2\pi)^{-m}\big(\,\widehat{u}, \widehat{v}\,\big). \tag{B.2.3b}$$

For $t \neq 0$ and $u \in \mathcal{S}\big(\mathbb{R}^m\big)$, we set

$$R_t u(\boldsymbol{x}) := u(t\boldsymbol{x}).$$

Then

$$\widehat{R_t u} = t^{-m} R_{t^{-1}} \widehat{u}. \tag{B.2.4}$$

If $u, v \in \mathcal{S}(\mathbb{R}^m)$, then their convolution

$$u * v(x) = \int_{\mathbb{R}^m} u(x-y)v(y)dy$$

is also a Schwartz function and

$$\mathscr{F}\big[\, u * v \,\big] = \mathscr{F}\big[\, u \,\big] \cdot \mathscr{F}\big[\, v \,\big]. \tag{B.2.5a}$$

$$\mathscr{F}\big[\, uv \,\big] = (2\pi)^{-m} \mathscr{F}\big[\, u \,\big] * \mathscr{F}\big[\, v \,\big]. \tag{B.2.5b}$$

Denote by $\mathcal{S}'(\mathbb{R}^m)$ the space of linear functionals $\mathcal{S}(\mathbb{R}^m) \to \mathbb{C}$ that are continuous with respect to the Fréchet structure on $\mathcal{S}(\mathbb{R}^m)$. We will refer to the elements of $\mathcal{S}'(\mathbb{R}^m)$ as *tempered distributions* on $\mathbb{R}^m$, and we will denote by $\langle -, - \rangle$ the natural pairing

$$\langle -, - \rangle : \mathcal{S}'(\mathbb{R}^m) \times \mathcal{S}(\mathbb{R}^m) \to \mathbb{C}, \;\; \langle \phi, u \rangle = \phi(u).$$

Note that we have an inclusion

$$C_b^0(\mathbb{R}^m) \hookrightarrow \mathcal{S}'(\mathbb{R}^m), \;\; u \mapsto L_u, \;\; \langle L_u, v \rangle = \langle u, v \rangle = \int_{\mathbb{R}^m} u(x)v(x)dx.$$

We can extend the Fourier transform to a map $\mathscr{F} : \mathcal{S}'(\mathbb{R}^m) \to \mathcal{S}'(\mathbb{R}^m)$ by setting

$$\big\langle \mathscr{F}\big[\, \phi \,\big], u \big\rangle := \big\langle \phi, \mathscr{F}\big[\, u \,\big] \big\rangle.$$

For example, the Dirac distribution δ_0 is a tempered distribution. Then

$$\big\langle \widehat{\delta}_0, u \big\rangle = \big\langle \delta_0, \widehat{u} \big\rangle = \widehat{u}(0) = \int_{\mathbb{R}^m} u(x)dx.$$

Thus the Fourier transform of δ_0 is the Lebesgue measure $\boldsymbol{\lambda}$. The Fourier inversion formula shows that

$$\widehat{\boldsymbol{\lambda}} = (2\pi)^m \delta_0.$$

Recall that a locally convex topological vector space is called *Montel* or *perfect* if every closed and bounded subset is compact. The space $\mathcal{S}\big(\mathbb{R}^m\big)$ is Montel; see [**67**, Sec. I.3] or [**154**, Sec. 34.4]. As discussed in Section 1.1.4, there are three remarkable topologies on $\mathcal{S}'\big(\mathbb{R}^m\big)$: the weak*, the Mackey and the strong topology. In the dual of a Montel space any weakly* convergent *sequence*[1] is also strongly convergent.

Example B.2.1 (Fourier transform of radial functions)**.** Suppose that we have a continuous function $f : [0,]\infty) \to \mathbb{C}$ such that the function

$$u : \mathbb{R}^m \to \mathbb{R}, \;\; u(\boldsymbol{x}) = f\big(\, |\boldsymbol{x}| \,\big).$$

is integrable. Then

$$\widehat{u}(\boldsymbol{\xi}) = \int_{\mathbb{R}^m} e^{\boldsymbol{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle} f\big(\, \boldsymbol{x}| \,\big) = \int_0^\infty \underbrace{\left(\int_{S^{m-1}} e^{\boldsymbol{i}\langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle} V_{S^{m-1}}\big[d\boldsymbol{\theta}\big] \right)}_{=:\Phi_m(r|\boldsymbol{\xi}|)} f(r) r^{m-1} dr$$

[1]We want to emphasize that this is a statement strictly about sequences, *not about generalized sequences.*

Using (A.1.11) we deduce that

$$\Phi_m(\rho) = \boldsymbol{\sigma}_{m-2}\int_{-1}^{1} e^{\boldsymbol{i}\rho t}\big(\,1-t^2\,\big)^{(m-3)/2}dt = \boldsymbol{\sigma}_{m-2}\int_{-1}^{1}\big(\,1-t^2\,\big)^{(m-3)/2}\cos(\rho t)dt$$

The last integral can be expressed in terms of Bessel's functions of the first kind. More precisely, for $\nu > -1/2$ we have (see [**86**, Eq. (5.10.3)])

$$J_\nu(\rho) = \frac{(\rho/2)^\nu}{\Gamma(1/2)\Gamma(\nu+1/2)}\int_{-1}^{1}\big(\,1-t^2\,\big)^{\nu-1/2}\cos(\rho t)dt$$

Hence, with $\nu = m/2-1$ we have

$$\Psi_m(\rho) = \boldsymbol{\sigma}_{m-2}\frac{\Gamma(1/2)\Gamma((m-1)/2)}{(\rho/2)^{m/2-1}}J_{m/2-1}(\rho)$$

$$= \frac{2\pi^{(m-1)/2}}{\Gamma((m-1)/2)}\frac{\Gamma(1/2)\Gamma((m-1)/2)}{(\rho/2)^{m/2-1}}J_{m/2-1}(\rho) = \frac{(2\pi)^{m/2}}{\rho^{m/2-1}}J_{m/2-1}(\rho).$$

Hence, the Fourier transform of $u((\boldsymbol{x}) = f(|\boldsymbol{x}|)$ is

$$\widehat{u}(\boldsymbol{\xi}) = (2\pi)^{m/2}|\boldsymbol{\xi}|^{1-m/2}\int_0^\infty J_{m/2-1}\big(\,r|\boldsymbol{\xi}|\,\big)r^{m/2}dr.$$

□

In this book we use frequently the *Poisson summation formula*

$$\forall u\in\mathcal{S}\big(\,\mathbb{R}^m\,\big),\ \ \forall a>0:\ \sum_{\vec{\ell}\in\mathbb{Z}^m} u\Big(\,\frac{2\pi}{a}\vec{\ell}\,\Big) = \Big(\,\frac{a}{2\pi}\,\Big)^m\sum_{\vec{k}\in\mathbb{Z}^m}\widehat{u}\big(\,a\vec{k}\,\big) \tag{B.2.6}$$

For a proof we refer to [**74**, §7.2].

For any open subset $\mathcal{O}\subset\mathbb{R}^m$ we denote by $C^{-\infty}\big(\,\mathcal{O}\,\big)$ the space of *generalized functions* or *distributions* on $\mathcal{O}$, i.e., continuous linear functionals (see [**74**, Sec. 2.1] for details)

$$C^\infty_{\mathrm{cpt}}\big(\,\mathcal{O}\,\big)\to\mathbb{C}.$$

Recall that a locally integrable function $f:\mathbb{R}^m\to\mathbb{R}$ is said to be homogeneous of degree $a\in\mathbb{R}$ if

$$f\big(\,t\boldsymbol{x}\,\big) = t^a f\big(\,\boldsymbol{x}\,\big),\ \ \forall t>0,\ \ \boldsymbol{x}\in\mathbb{R}^m\setminus 0.$$

In general. a distribution $u\in C^{-\infty}\big(\,\mathbb{R}^m\setminus 0\,\big)$ is homogeneous of degree a iff for any $t>0$ we have

$$\langle u, tR_t\varphi\rangle = t^{-a}\langle u,\varphi\rangle,\ \ \forall\varphi\in C^\infty_{\mathrm{cpt}}\big(\,\mathbb{R}^m\setminus 0\,\big).$$

Informally, this reads

$$\int_{\mathbb{R}^m} u(\boldsymbol{x})t\varphi(t\boldsymbol{x})d\boldsymbol{x} = t^{-a}\int_{\mathbb{R}^m} u(\boldsymbol{x})\varphi(\boldsymbol{y})d\boldsymbol{x}.$$

A temperate distribution $u\in\mathcal{S}'\big(\,\mathbb{R}^m\,\big)$ is homogeneous of degree a if its restriction on $\mathbb{R}^m\setminus 0$ is homogeneous of degree a. The equality (B.2.4) show that $u\in\mathcal{S}'\big(\,\mathbb{R}^m\,\big)$ is homogeneous of degree a iff its Fourier transform is homogeneous of degree $-m-a$.

B.3. Basic facts about spherical harmonics

We survey here a few classical facts about spherical harmonics that we needed in the main body of the paper. For proofs and more details we refer to our main source, [105].

We denote by $\mathcal{H}_{n,d}$ the space of homogeneous, harmonic polynomials of degree n in d variables. We regard such polynomials as functions on $\mathbb{R}^d$, and we denote by $\mathcal{Y}_{n,d}$ the subspace of $C^\infty(S^{d-1})$ spanned by the restrictions of these polynomials to the unit sphere. We have

$$\dim \mathcal{H}_{n,d} = \dim \mathcal{Y}_{n,d} = M(n,d) = \binom{d+n-1}{n} - \binom{d+n-3}{n-2}$$
$$= \frac{2n+d-2}{n+d-2}\binom{n+d-2}{d-2} \sim 2\frac{n^{d-2}}{(d-2)!} \text{ as } n \to \infty.$$

Observe that

$$M(0,d) = 1, \;\; M(1,d) = d, \;\; M(2,d) = \binom{d+1}{2} - 1. \tag{B.3.1}$$

The space $\mathcal{Y}_{n,d}$ is the eigenspace of the Laplace operator on S^{d-1} corresponding to the eigenvalue $\lambda_n(d) = n(n+d-2)$.

The *Legendre polynomial* $P_{n,d}(t)$ *of degree* n *and order* d is given by the Rodriguez formula

$$P_{n,d}(t) = (-1)^n R_n(d)(1-t^2)^{-\frac{d-3}{2}}\left(\frac{d}{dt}\right)^n (1-t^2)^{n+\frac{d-3}{2}}, \tag{B.3.2}$$

where $R_n(d)$ is the Rodriguez constant

$$R_n(d) = 2^{-n}\frac{\Gamma(\frac{d-1}{2})}{\Gamma(n+\frac{d-1}{2})} = 2^{-n}\frac{1}{\left(n+\frac{d-3}{2}\right)_n},$$

where we recall that $(x)_k := x(x-1)\cdots(x-k+1)$.

Equivalently, they can be defined recursively via the relations

$$P_{0,d}(t) = 1, \;\; P_{1,d}(t) = t,$$
$$(n+d-2)P_{n+1,d}(t) - (2n+d-2)tP_{n,d}(t) + nP_{n-1,d}(t) = 0, \;\; n > 0.$$

In particular, this shows that

$$P_{2,d}(t) = \frac{1}{d-1}\left(dt^2 - 1\right).$$

The Legendre polynomials are normalized by the equality

$$P_{n,d}(1) = 1, \;\; \forall d \geq 2, \;\; n \geq 0.$$

More generally, for any $n > 0$, $d \geq 2$, and any $0 < j \leq n$, we have

$$P^{(j)}_{n,d}(1) = (-1)^n R_n(d)\binom{n+j}{j}\left\{\frac{D_t^n(1-t)^{n+\frac{d-3}{2}}}{(1-t)^{\frac{d-3}{2}}}\cdot\frac{D_t^j(1+t)^{n+\frac{d-3}{2}}}{(1+t)^{\frac{d-3}{2}}}\right\}_{t=1}, \;\; D_t := \frac{d}{dt},$$
$$= 2^{n-j}R_n(d)\binom{n+j}{j}\left(n+\frac{d-3}{2}\right)_n\cdot\left(n+\frac{d-3}{2}\right)_j,$$

which implies

$$P^{(j)}_{n,d}(1) = 2^{-j}\binom{n+j}{j}\left(n+\frac{d-3}{2}\right)_j. \tag{B.3.3}$$

Fix $y \in S^{d-1}$. Denote by $\bullet$ the canonical inner product on $\mathbb{R}^d$. Then the function

$$x \mapsto P_{n,d}(x \bullet y)$$

belongs to the eigenspace $\mathcal{H}_{n,d}$. Note that $\arccos(x \bullet y)$ is the geodesic distance between $x, y \in S^{d-1}$.

If $(\Psi_k)_{1\leq k\leq M(s,d)}$ is an orthonormal basis of $\mathcal{H}_{n,d}$, then we have the *addition formula*

$$\forall x, y \in S^{d-1}, \quad \sum_{k=1}^{M(n,d)} \Psi_k(x)\Psi_k(y) = \frac{M(n,d)}{\boldsymbol{\sigma}_{d-1}} P_{n,d}(x \bullet y) \tag{B.3.4}$$

where $\boldsymbol{\sigma}_{d-1}$ denotes the "area" of the unit sphere in $\mathbb{R}^d$.

Denote by $\boldsymbol{P}_n = \boldsymbol{P}_{n,d}$ the ortogonal projection $L^2\big(S^{d-1}\big) \to \mathcal{H}_{n,d}$. Observe that

$$\mathcal{K}_n(x,y) = \mathcal{K}_{n,d}(x,y) = \sum_{k=1}^{M(n,d)} \Psi_k(x)\Psi_k(y)$$

is the integral kernel of the operator $\boldsymbol{P}_n$, i.e., $\forall f \in L^2(S^{d-1})$

$$\boldsymbol{P}_{n,d} f(x) = \int_{S^{d-1}} \mathcal{K}_{n,d}(x,y) \operatorname{vol}_{S^{d-1}}\big[\,dy\,\big].$$

Theorem B.3.1 (Funk-Hecke formula)**.** *Let $\Psi \in \mathcal{H}_{n,d}$ and $f \in C^0\big([-1,1]\big)$. Then for any $x \in S^{d-1}$ we have*

$$\int_{S^{d-1}} f(x \bullet y)\Psi(y) \operatorname{vol} S^{d-1}\big[\,dy\,\big] = \lambda_n \Psi(x), \tag{B.3.5}$$

where

$$\lambda_n = \boldsymbol{\sigma}_{d-1} \int_{-1}^{1} f(t) P_{n,d}(t)\big(1-t^2\big)^{\frac{d-3}{2}} dt.$$

□

We want to describe an inductive construction of an orthonormal basis of $\mathcal{Y}_{n,d}$. We start with the case $d=2$. For any $m \in \mathbb{Z}$, we set

$$\varphi_m(\theta) = \begin{cases} \cos(m\theta), & m \leq 0 \\ \sin(m\theta), & m > 0. \end{cases}, \quad t_m = \|\varphi_m\|_{L^2} = \begin{cases} (2\pi)^{1/2}, & m = 0 \\ \pi^{1/2}, & m > 0. \end{cases}, \quad \Phi_m = \frac{1}{t_m}\varphi_m.$$

Then $\mathcal{B}_{0,2} = \{\Phi_0\}$ is an orthonormal basis of $\mathcal{Y}_{0,2}$, while $\mathcal{B}_{n,2} = \{\Phi_{-n}, \Phi_n\}$ is an orthonormal basis of $\mathcal{Y}_{n,2}$, $n > 0$.

Assuming now that we have produced orthonormal bases $\mathcal{B}_{n,d-1}$ of all the spaces $\mathcal{Y}_{n,d-1}$, we indicate how to produce orthonormal bases in the harmonic spaces $\mathcal{Y}_{n,d}$. This requires the introduction of the Legendre polynomials and their associated functions.

For any $d \geq 3$, $n \geq 0$ and $0 \leq j \leq n$, we define the *normalized associated Legendre functions*

$$\widehat{P}^j_{n,d}(t) := C_{n,j,d}(1-t^2)^{j/2} P^{(j)}_{n,d}(t),$$

where

$$C_{n,j,d} := \frac{[n+d-3]_{d-3}}{\Gamma(\frac{d-1}{2})} \left(\frac{(2n+d-2)}{2^{d-2}[n+d+j-3]_{2j+d-3}} \right)^{1/2}. \tag{B.3.6}$$

When $d = 3$, the above formulæ take the form

$$\widehat{P}_{n,j,3}(t) = \sqrt{\frac{(n+\frac{1}{2})(n-j)!}{(n+j)!}}(1-t^2)^j P_{n,3}^{(j)}(t). \tag{B.3.7}$$

For any $0 \leq j \leq n$, and any $d > 2$ we define a linear map

$$\mathcal{T}_{n,j,d} : \mathcal{Y}_{j,d-1} \to \mathcal{Y}_{n,d}, \quad Y \mapsto \mathcal{T}_{n,j,d}[Y],$$

$$\mathcal{T}_{n,j,d}[Y](\boldsymbol{x}) = \widehat{P}_{n,d}^{j}(x_d) \cdot Y\left(\frac{1}{\|\boldsymbol{x}'\|}\boldsymbol{x}'\right),$$

$$\forall \boldsymbol{x} = (\boldsymbol{x}', x_d) \in S^{d-1}, \ \boldsymbol{x}' = (x_1, \dots, x_{d-1}) \neq 0.$$

Note that for $\boldsymbol{x} = (\boldsymbol{x}', x_d) \in S^{d-1}$ we have

$$\|\boldsymbol{x}'\| = (1-x_d)^{1/2} \ \text{ and } \ \widehat{P}_{n,d}^{j}(x_d) = C_{n,j,d}(1-x_d^2)^{j/2}P_{n,d}^{(j)}(x_d) = C_{n,j,d}\|\boldsymbol{x}'\|^j P_{n,d}^{(j)}(x_d),$$

so that

$$\mathcal{T}_{n,j,d}[Y](\boldsymbol{x}) = C_{n,j,d}P_{n,d}^{(j)}(x_d)\widetilde{Y}(\boldsymbol{x}'), \ \forall \boldsymbol{x} = (\boldsymbol{x}', x_d) \in S^{d-1},$$

where $\widetilde{Y}$ denotes the extension of Y as a homogeneous polynomial of degree j in $(d-1)$-variables. The sets $\mathcal{T}_{n,j,d}[\mathcal{B}_{j,d-1}]$, $0 \leq j \leq n$ are disjoint, and their union is an orthonormal basis of $\mathcal{Y}_{n,d}$ that we denote by $\mathcal{B}_{n,d}$.

The space $\mathcal{Y}_{0,d}$ consists only of constant functions and $\mathcal{B}_{0,d} = \{\, \boldsymbol{\sigma}_{d-1}^{-\frac{1}{2}} \,\}$. The orthonormal basis $\mathcal{B}_{1,d}$ of $\mathcal{Y}_{1,d}$ obtained via the above inductive process is

$$\mathcal{B}_{1,d} = \left\{\, C_0 x_i, \ \ 1 \leq i \leq d \,\right\} = \left\{\boldsymbol{\sigma}_{d-2}^{-\frac{1}{2}} C_{1,0,d} x_i; \ \ 1 \leq i \leq d\,\right\}. \tag{B.3.8}$$

The orthonormal basis $\mathcal{B}_{2,d}$ of $\mathcal{Y}_{2,d}$ is

$$C_1(dx_i^2 - r^2), \ \ 1 \leq i < d, \ \ C_2 x_i x_j, \ \ 1 \leq i < j \leq d, \tag{B.3.9}$$

where $r^2 = x_1^2 + \cdots + x_d^2$, and the positive constants C_0, C_1, C_2 are found from the equalities

$$C_0^2 \int_{S^{d-1}} x_1^2 \, |dS(\boldsymbol{x})| = C_1^2 \int_{S^{d-1}} (d^2 x_1^4 - 2dx_1^2 + 1)\, |dS(\boldsymbol{x})| = C_2^2 \int_{S^{d-1}} x_1^2 x_2^2 |dS(\boldsymbol{x})| = 1,$$

aided by the classical identities, (2.3.13)

$$\int_{S^{d-1}} x_1^{2h_1} \cdots x_d^{2h_d} \, |dS(\boldsymbol{x})| = \frac{2\Gamma(\frac{2h_1+1}{2}) \cdots \Gamma(\frac{2h_d+1}{2})}{\Gamma(\frac{2h+d}{2})}, \ \ h = h_1 + \cdots + h_d. \tag{B.3.10}$$

B.4. Some asymptotic estimates

We want to discuss the large m asymptotics of

$$\frac{h_m(\mathfrak{a})}{d_m(\mathfrak{a})} = \frac{2I_{m+3}(\mathfrak{a})}{(m+2)I_{m+1}(\mathfrak{a})}, \ \ q_m(\mathfrak{a}) = \frac{s_m(\mathfrak{a})h_m(\mathfrak{a})}{d_m(\mathfrak{a})^2} = \frac{m}{m+2} \cdot \underbrace{\frac{I_{m-1}(\mathfrak{a})I_{m+3}(\mathfrak{a})}{I_{m+1}(\mathfrak{a})^2}}_{=:R_m(\mathfrak{a})},$$

for various choices of amplitudes $\mathfrak{a}$. Set $w_{\mathfrak{a}}(t) := \mathfrak{a}(t)^2$. Recall (2.3.27)

$$\begin{aligned} C_m(\mathfrak{a}) &\sim 2^{5/2}\left(\frac{h_m(\mathfrak{a})}{d_m(\mathfrak{a})}\right)^{m/2}\Gamma\left(\frac{m+3}{2}\right)m^{-1/2} \\ &\sim 2^{\frac{m+5}{2}}\left(\frac{I_{m+3}(\mathfrak{a})}{(m+2)I_{m+1}(\mathfrak{a})}\right)^{m/2}\Gamma\left(\frac{m+3}{2}\right)m^{-1/2}\ \text{ as } m\to\infty. \end{aligned} \tag{B.4.1}$$

Example B.4.1. Suppose that $w_{\mathfrak{a}}(t) = e^{-t^2}$. In this case

$$I_k(\mathfrak{a}) = \int_0^\infty t^k e^{-t^2}dt = \frac{1}{2}\int_0^\infty s^{\frac{k-1}{2}}e^{-s}ds = \frac{1}{2}\Gamma\left(\frac{k+1}{2}\right).$$

Hence $I_{k+2} = \frac{k+1}{2}I_k$, $\forall k$,

$$\frac{I_{m+3}(\mathfrak{a})}{I_{m+1}(\mathfrak{a})} = \frac{m+2}{2},\quad \frac{h_m(\mathfrak{a})}{d_m(\mathfrak{a})} = 1,\quad q_m = \frac{m(m+4)}{(m+2)^2} < 1,\quad \forall m.$$

We deduce

$$C_m(w) \overset{(2.3.27)}{\sim} \frac{2^{\frac{5}{2}}}{\sqrt{m}\pi^{\frac{m+1}{2}}}\Gamma\left(\frac{m+3}{2}\right)\ \text{ as } m\to\infty,$$

and Stirling's formula implies

$$\log C_m(w) \sim \frac{m}{2}\log m\ \text{ as } m\to\infty. \tag{B.4.2}$$

□

Example B.4.2. Suppose that

$$w_{\mathfrak{a}}(t) = \exp\bigl(-(\log t)\log(\log t)\bigr),\quad \forall t\geq 1.$$

Observe that

$$I_k(\mathfrak{a}) = \int_0^1 r^k w(r)dr + \int_1^\infty r^k\exp\bigl(-(\log r)\log(\log r)\bigr)dr.$$

This proves that

$$I_k(\mathfrak{a}) \sim J_k := \int_1^\infty r^k\exp\bigl(-(\log r)\log(\log r)\bigr)dr\ \text{ as } k\to\infty.$$

Using the substitution $r = e^t$ we deduce

$$J_k = \int_0^\infty e^{(k+1)t - t\log t}dt.$$

We want to investigate the large λ asymptotics of the integral

$$T_\lambda = \int_0^\infty e^{-\phi_\lambda(t)}dt,\quad \phi_\lambda(t) = \lambda t - t\log t. \tag{B.4.3}$$

We will achieve this by relying on the Laplace method [**29**, Chap. 4]. Note that

$$\phi_\lambda'(t) = \lambda - \log t - 1,\quad \phi_\lambda''(t) = -\frac{1}{t}.$$

Thus $\phi_\lambda(t)$ has a unique critical point

$$\tau = \tau(\lambda) := e^{\lambda-1}.$$

We make the change in variables $t = \tau s$ in (B.4.3). Observe that

$$\lambda e^{\lambda-1}s - e^{\lambda-1}s\log(e^{\lambda-1}s) = e^{\lambda-1}s - (\lambda-1)e^{\lambda-1}s - e^{\lambda-1}\log s = e^{\lambda-1}s(1-\log s)$$

and we deduce

$$T_\lambda = \tau\int_0^\infty e^{-\tau h(s)}ds, \;\; h(s) = s(\log s - 1).$$

The asymptotics of the last integral can be determined using the Laplace method and we have, [29, §4.1]

$$T_\lambda \sim \tau e^{-\tau h(1)}\sqrt{\frac{2\pi}{\tau h''(1)}} = \sqrt{2\pi\tau}e^\tau.$$

Hence

$$J_k = T_{k+1} \sim \sqrt{2\pi\tau(k+1)}e^{\tau(k+1)} = \sqrt{2\pi e^k}e^{e^k} \;\text{ as } k\to\infty.$$

In this case

$$\lim_{m\to\infty} q_m(\mathfrak{a}) = \infty,$$

and

$$\log\left(\frac{h_m}{d_m}\right) \sim e^{m+4} - e^{m+2} \;\text{ as } m\to\infty.$$

Hence

$$\log C_m(\mathfrak{a}) \sim \frac{m}{2}e^{m+2}(e^2-1) \;\text{ as } m\to\infty. \tag{B.4.4}$$

□

Example B.4.3. Fix $C > 0$ and $\alpha > 1$. Suppose that

$$w_{\mathfrak{a}}(t) = \exp\bigl(-C(\log t)^\alpha\bigr), \forall t > 1.$$

Arguing as in Example B.4.2 we deduce that as $k\to\infty$

$$I_k(\mathfrak{a}) \sim \int_1^\infty t^k\exp\bigl(-C(\log t)^\alpha\bigr)dr = \int_0^\infty e^{(k+1)t - Ct^\alpha}dt.$$

Again, set

$$T_\lambda := \int_0^\infty e^{-\phi_\lambda(t)}dt, \;\; \phi_\lambda(t) := Ct^\alpha - \lambda t.$$

We determine the asymptotics of T_λ as $\lambda\to\infty$ using the Laplace method. Note that

$$\phi'_\lambda(t) = \alpha C t^{\alpha-1} - \lambda.$$

The function ϕ_λ has a unique critical point

$$\tau = \tau(\lambda) = \left(\frac{\lambda}{\alpha C}\right)^{\frac{1}{\alpha-1}}.$$

Observe that

$$\phi_\lambda(\tau s) = a(s^\alpha - bs), \;\; a := \left(\frac{\lambda}{C^{1/\alpha}\alpha}\right)^{\frac{\alpha}{\alpha-1}}, \;\; b := \alpha^{\frac{1}{\alpha-1}},$$

$$T_\lambda = \tau(\lambda)\int_0^\infty e^{-a(s^\alpha - bs)}ds.$$

We set $g(s) := s^\alpha - bs$. Using the Laplace method [29, §4.2] we deduce

$$T_\lambda \sim \tau(\lambda)e^{-ag(1)}\sqrt{\frac{2\pi}{ag''(1)}} = \sqrt{\frac{2\pi}{a\alpha(\alpha-1)}}e^{a(b-1)}.$$

Hence

$$\begin{aligned}\log T_\lambda &\sim \left(\frac{\lambda^\alpha}{C}\right)^{\frac{1}{\alpha-1}} \frac{\alpha^{\frac{1}{\alpha-1}}-1}{\alpha^{\frac{\alpha}{\alpha-1}}} =: Z(\alpha,C)\lambda^{\frac{\alpha}{\alpha-1}},\\ \log I_{m+3}(\mathfrak{a}) - \log I_{m+1}(\mathfrak{a}) &\sim Z(\alpha,C)m^{\frac{\alpha}{\alpha-1}}\left(\left(1+\frac{4}{m}\right)^{\frac{\alpha}{\alpha-1}} - \left(1+\frac{2}{m}\right)^{\frac{\alpha}{\alpha-1}}\right)\\ &\sim \frac{2Z(\alpha,C)}{\alpha-1}m^{\frac{1}{\alpha-1}},\ \ m\to\infty,\end{aligned}$$

so that

$$\log C_m(\mathfrak{a}) \sim \frac{Z(\alpha,C)}{\alpha-1}m^{\frac{\alpha}{\alpha-1}},\ \ m\to\infty. \tag{B.4.5}$$

Similarly

$$\begin{aligned}\log R_m(w) &\sim \log T_m + \log T_{m+4} - 2\log T_{m+2}\\ &\sim Z(\alpha,C)\Big(m^{\frac{\alpha}{\alpha-1}} + (m+4)^{\frac{\alpha}{\alpha-1}} - 2(m+2)^{\frac{\alpha}{\alpha-1}}\Big)\\ &= Z(\alpha,C)m^{\frac{\alpha}{\alpha-1}}\left(1+\left(1+\frac{4}{m}\right)^{\frac{\alpha}{\alpha-1}} - 2\left(1+\frac{2}{m}\right)^{\frac{\alpha}{\alpha-1}}\right)\\ &\sim Z(\alpha,C)m^{\frac{\alpha}{\alpha-1}} \times \frac{8}{m^2} \times \frac{\alpha}{\alpha-1}\left(\frac{\alpha}{\alpha-1}-1\right) = \frac{8\alpha Z(\alpha)}{(\alpha-1)^2}m^{\frac{2-\alpha}{\alpha-1}}.\end{aligned}$$

Hence

$$\lim_{m\to\infty} R_m = \times\begin{cases}\infty, & \alpha<2,\\ e^{16Z(2,C)}, & \alpha=2,\\ 1, & \alpha>2.\end{cases} \tag{B.4.6}$$

□

B.5. Reproducing Hilbert Kernel Spaces

In the more than a century since their appearance on the mathematical scene the Reproducing Kernel Hilbert Spaces have found applications in diverse areas: complex analysis, numerical analysis, quantum mechanics, Gaussian processes and machine learning, to name a few. The goal of this section is to survey, mostly without proofs, some basic properties of such spaces. Our main sources of inspiration are [9, 129] to which we refer for proofs and more details.

Let X be a set. Recall that $\mathbb{R}^X$ is the space of functions $X\to\mathbb{R}$. For every $x\in X$ we denote by $\mathbf{Ev}_x$ the evaluation at x, i.e. the linear map

$$\mathbf{Ev}_x : \mathbb{R}^X \to \mathbb{R},\ \ \mathbf{Ev}_x\big[\, f\,\big] = f(x).$$

A (real) *Reproducing Kernel Hilbert Space* over X, or RKHS henceforth, is a *vector subspace* $\mathcal{H}\subset\mathbb{R}^X$ with the following properties

(i) It is equipped with an inner product $(-,-)_{\mathcal{H}}$ making it into a real Hilbert space.

(ii) For every $x\in X$ the linear functional $\mathbf{Ev}_x : \mathcal{H}\to\mathbb{R}$ is continuous with respect to the Hilbert norm.

From the Riez representation theorem and (ii) we deduce that for every $x \in X$ there exists $K_x \in \mathcal{H}$ such that

$$\big(K_x, h \big)_{\mathcal{H}} = \mathbf{Ev}_x \big[h \big] = h(x), \;\; \forall h \in \mathcal{H}.$$

The resulting function

$$K : X \times X \to \mathbb{R}, \;\; K(x,y) = K_x(y) = \mathbf{Ev}_y \big[K_x \big] = (K_y, K_x)_{\mathcal{H}},$$

is called the *reproducing kernel* or the *kernel* of the RKHS $\mathcal{H}$.

There is a natural map $\Phi : X \to \mathcal{H}$, $\Phi(x) = K_x$. Note that

$$K(x,y) = \big(\Phi(x), \Phi(y) \big)_{\mathcal{H}}, \;\; \forall x, y \in X.$$

In machine learning this map is known as the *feature map*.

Note that if X is a topological space and K is continuous, then the feature map is continuous as a map from X to the Hilbert space $\mathcal{H}$. Indeed, for any $x_0 \in X$, the function $u : X \to \mathbb{R}$

$$u(x) = \|\Phi(x) - \Phi(x_0)\|^2_{\mathcal{H}} = K(x,x) - 2K(x,x_0) + K(x_0,x_0)$$

and

$$\lim_{x \to x_0} u(x) = 0.$$

Example B.5.1. The feature map is a disguised version of a standard geometric construction. More precisely, given a set X and a finite dimensional vector space V of real valued functions on X, we get a tautological map

$$\mathbf{Ev} : X \to V^*, \;\; x \mapsto \mathbf{Ev}_x \, .$$

The map $\mathbf{Ev}$ is injective if and only if the vector space V *separates the points*, i.e., $\forall x, y \in X$, $\exists v \in V$ such that $v(x) \neq v(y)$.

Fix an inner product $(-,-)$ on V. Since V is finite dimensional, the evaluation maps $\mathbf{Ev}_x : V \to \mathbb{R}$ are continuous with respect to this inner product for any $x \in X$. The pair $\big(V, (-,-) \big)$ is an RKHS.

The inner product induces a dual inner product $(-,-)_{V^*}$ on V^*, we can identify V^* with V and the evaluation map $\mathbf{Ev} : X \to V^* \cong V$ is the feature map. The reproducing kernel is

$$K(x,y) = \big(\, \mathbf{Ev}_x, \mathbf{Ev}_y \, \big)_{V^*}.$$

Suppose that X is a subset of a finite dimensional Euclidean space V. The inner product on V induces a duality isomorphism

$$V \ni \boldsymbol{v} \mapsto \boldsymbol{v}_{\downarrow} \in V^*, \;\; \boldsymbol{v}^{\downarrow}(\boldsymbol{u}) = (\boldsymbol{v}, \boldsymbol{u}), \;\; \forall \boldsymbol{u} \in V.$$

In particular, we get a map

$$X \ni \boldsymbol{x} \mapsto \boldsymbol{x}^{\downarrow} \in V^*.$$

Assume for simplicity that X is not contained in any proper subspace of V. Then the map

$$V^* \ni \xi \mapsto \xi\big|_X \in \mathbb{R}^X$$

is one-to-one and we denote by $\mathcal{H}$ the image of this map. The inner product on V^* induces an inner product on $\mathcal{H}$,

$$\big(\boldsymbol{v}_0^{\downarrow}\big|_X, \boldsymbol{v}_1^{\downarrow}\big|_X \big)_{\mathcal{H}} := \big(\boldsymbol{v}_0^{\downarrow}, \boldsymbol{v}_1^{\downarrow} \big)_{V^*} = \big(\boldsymbol{v}_0, \boldsymbol{v}_1 \big)_V.$$

Then $\mathcal{H}$ is a RKHS with kernel $K(x, y) = (x, y)_V$ and feature map

$$X \mapsto \mathcal{H}, \;\; \boldsymbol{x} \mapsto \boldsymbol{x}^{\downarrow}\big|_X \in \mathcal{H}.$$

Note that the function $\mathbf{Ev}_{\boldsymbol{x}} : \mathcal{H} \to \mathbb{R}$ coincides with $\boldsymbol{x}^{\downarrow}\big|_{\boldsymbol{X}}$. □

Observe that the reproducing kernel of a RKHS is a symmetric function, i.e.,

$$K(x, y) = K(y, x) = \big(K_x, K_y \big)_{\mathcal{H}}, \;\; \forall x, y \in X. \qquad (\mathbf{RK}_1)$$

Indeed,

$$K(x, y) = K_x(y) = \mathbf{Ev}_y(K_x) = \big(K_y, K_x \big)_{\mathcal{H}}.$$

For any $x_1, \dots, x_n \in X$ we denote by $\mathbb{G}_K(x_1, \dots, x_n)$ the symmetric $n \times n$ matrix

$$\mathbb{G}_K(x_1, \dots, x_n) := \big(K(x_i, x_j) \big)_{1 \leq i,j \leq n}.$$

Observe that

$$\mathbb{G}_K(x_1, \dots, x_n) \geq 0, \;\; \forall n, \; \forall x_1, \dots, x_n \in X. \qquad (\mathbf{RK}_2)$$

Indeed, $\mathbb{G}_K(x_1, \dots, x_n)$ is Grammian of the functions K_{x^i}

$$\mathbb{G}_K(x_1, \dots, x_n) = \big((K_{x_i}, K_{x_j})_{\mathcal{H}} \big)_{1 \leq i,j \leq n},$$

and the Grammians are *positive semidefinite* matrices, i.e., all their eigenvalues are *nonnegative*.

The rank of $\mathbb{G}_K(x_1, \dots, x_n)$ is the dimension of the space span $\big\{ K_{x_1}, \dots, K_{x_n} \big\}$. In particular, we deduce that if $\det \mathbb{G}_K(x_1, x_2) \neq 0$ then $K_{x_1} \neq K_{x_2}$. We have the following consequence.

Corollary B.5.2. *If the reproducing kernel K of an RKHS $\mathcal{H}$ over X satisfies*

$$\det \mathbb{G}_K(x_1, x_2) \neq 0, \;\; \forall x_1, x_2 \in X, \;\; x_1 \neq x_2, \qquad (B.5.1)$$

then the feature map $\Phi : X \to \mathcal{H}$ is injective. □

Definition B.5.3. We define a *(reproducing) kernel* on a *topological space* X to be a continuous symmetric function $K : X \times X \to \mathbb{R}$ satisfying ($\mathbf{RK}_2$). A *reproducing kernel* on a *set* X is a reproducing kernel on X equipped with the discrete topology.

We denote by $\mathcal{K}(X)$ the set of kernels on X. We denote by $\mathcal{K}^+(X)$ the collection of kernels K such that for any distinct points $x_1, \dots, x_n \in X$ the symmetric matrix $\mathbb{G}_K(x_1, \dots, x_n)$ is *positive definite*, i.e., all its eigenvalues are *positive*. □

Theorem B.5.4. *Let X be a set.*

(i) *The set of reproducing kernels $\mathcal{K}(X)$ is a convex cone in the vector space of functions $X \times X \to \mathbb{R}$.*

(ii) *If $(K_n)_{n \in \mathbb{N}}$ is a sequence of kernels on X that converges pointwisely to a function $K : X \times X \to \mathbb{R}$, then $K \in \mathcal{K}(X)$.*

(iii) *If $K_1, K_2 \in \mathcal{K}(X)$ then $K_1 \cdot K_2 \in \mathcal{K}(X)$.*

(iv) *If X' is another set, $K \in \mathcal{K}(X)$, $K' \in \mathcal{K}(X')$ and we define*

$$K \otimes K' : (X \times X') \times (X \times'), \;\; K \otimes K'\big((x, x'), (y, y') \big) = K(x, y) K'(x', y')$$

then $K \otimes K' \in \mathcal{K}(X \times X')$.

□

The only non-obvious part of the above result is (iii). It is a consequence of a less popular result of linear algebra stating that the Hadamard product of two positive semidefinite symmetric matrices is a positive semidefinite matrix; see [**129**, Sec. 4.2] . We recall that the Hadamard product of two matrices $A = (a_{ij})_{1\leq i,j\leq n}$, $B = (b_{ij})_{1\leq i,j\leq n}$ is the matrix

$$A * B = \big(a_{ij}b_{ij} \big)_{1\leq i,j\leq n}.$$

The following example of RKHS is in a certain sense universal.

Example B.5.5 (The RKHS of a Gaussian process)**.** We follow the presentation in [**69**, Sec. 2.61].

Let $(Z_x)_{x\in X}$ be a centered Gaussian process parametrized by a topological space X. We assume that the covariance kernel

$$C : X \times X \to \mathbb{R},\;\; C(x, y) = \mathbb{E}\big[\, Z_x Z_y \,\big]$$

is continuous, i.e., the map

$$X \ni x \to X_x \in L^2(\Omega, \mathcal{S}, \mathbb{P}),\;\; x \to Z_x$$

is continuous.

We denote by $\mathcal{Z}$ Gaussian Hilbert space determined by the Gaussian stochastic process $(Z_x)_{x\in X}$, i.e., the closure in $L^2\big(\mathcal{S}, \mathbb{P}\big)$ of the vector subspace

$$V := \operatorname{span}(Z_x)_{x\in X}.$$

We define

$$R : \mathcal{Z} \to \mathbb{R}^X,\;\; Z \mapsto R[Z],\;\; R[Z](x) = \mathbb{E}\big[\, Z Z_x \,\big],\; \forall x \in X.$$

Note that the function $R[Z] : X \to \mathbb{R}$ is continuous since the map

$$X \ni x \mapsto Z_x \in L^2$$

is continuous.

Observe also that the map R is injective. Indeed, if for some $Z_0 \in \mathcal{Z}$ we have $R[Z_0](x) = 0$, $\forall x \in X$, then $\mathbb{E}\big[\, Z_0 Z \,\big] = 0$, $\forall Z \in V$. Since V is dense in $\mathcal{Z}$ we deduce $\mathbb{E}\big[\, Z_0 Z \,\big] = 0$, $\forall Z \in \mathcal{Z}$, so $Z_0 = 0$.

We denote by $\mathcal{H}$ the image of R, $\mathcal{H} = R(\mathcal{Z}) \subset C(X) \subset \mathbb{R}^X$. The space $\mathcal{H}$ is a Hilbert space with respect to the inner product

$$\big\langle\, R[Z_0], R[Z_1] \,\big\rangle_{\mathcal{H}} := \mathbb{E}\big[\, Z_0 Z_1 \,\big],\;\; \forall R[Z_0], R[Z_1] \in \mathcal{H}.$$

The map R is an isomorphism of Hilbert spaces $\mathcal{Z} \to \mathcal{H}$.

Note that $R[Z_y] = C_y$, $C_y(x) = C(x, y) = C(y, x)$. Since the family (Z_y) is dense in $\mathcal{Z}$ we deduce that thefunctions C_y, $y \in X$, span a dense subspace of $\mathcal{H}$.

The map $\mathbf{Ev}_x : \mathcal{H} \to \mathbb{R}$, $R[Z] \mapsto R[Z](x)$ is continuous with respect to the inner product $\langle -, - \rangle_{\mathcal{H}}$ since

$$R[Z](x) := \mathbb{E}\big[\, Z Z_x \,\big] = \big\langle\, R[Z], R[Z_x] \,\big\rangle_{\mathcal{H}}.$$

We also have a map

$$\boldsymbol{F} = \boldsymbol{F}_C : X \to \mathcal{H},\;\; x \mapsto R[Z_x].$$

Observe that for any $x, y \in X$ we have

$$\big\langle R[Z_x], R[Z_y] \big\rangle = \mathbb{E}\big[Z_x Z_y \big] = C(x,y).$$

Thus, $\mathcal{H}$ is a RKHS and its reproducing kernel is the covariance kernel of the process $(Z_x)_{x\in X}$. Note that $\mathcal{H} \subset C(X)$. The feature map of $\mathcal{H}$ is $\boldsymbol{F}_C$. The space $\mathcal{H}$ is also known as the *Cameron-Martin space* of the Gaussian stochastic process $(Z_x)_{x\in X}$; see [**76**, Def. 8.14]. □

It turns out that Example B.5.5 is universal.

Theorem B.5.6 (Moore)**.** *Let X be a topological space. For any reproducing kernel $K \in \mathcal{K}(X)$, there exists a unique RKHS on $\mathcal{H} \subset C(X)$ whose reproducing kernel is K.*

Proof. *Existence.* Let $K \in \mathcal{K}(X)$. The Kolmogorov existence/consistency theorem [**118**, Sec. 1.5.2] shows that there exists a centered Gaussian process $(Z_x)_{x\in X}$ with covariance kernel K.

Uniqueness. Suppose that $\mathcal{H}_1, \mathcal{H}_2 \subset \mathbb{R}^X$ are two RKHS's with kernels K_1, K_2. We want to show that if $K_1 = K_2$, then $\mathcal{H}_1 = \mathcal{H}_2$. Set $K = K_1 = K_2$ We outline the main ideas referring for details to[**129**, Sec. 2.1].

For $i = 1, 2$ we denote by $(-,-)_i$ the inner product in $\mathcal{H}_i$ and by $\|-\|_i$, the corresponding norm. We set

$$V_K := \operatorname{span}\big\{ K_x;\ \ x \in X \big\}.$$

Note that $V_K \subset \mathcal{H}_i$ and

$$(u,v)_1 = (u,v)_2,\ \ \forall u, v \in V_K.$$

It is easy to see that V_K is dense in $\mathcal{H}_i$. Moreover, if $h_i \in \mathcal{H}_i$ and $\|v_n - h_i\|_i \to 0$, then $v_n(x) \to h_i(x)$, $\forall x \in X$. It suffices to show that if $h \in \mathcal{H}_1$, then $h \in H_2$.

Since $h \in \mathcal{H}_1$, there exists a sequence (v_n) in V_k such that $\|v_n - h\|_1 \to 0$. Hence (v_n) is Cauchy in $\mathcal{H}_1$. Since $\|v_n - v_m\|_1 = \|v_n - v_m\|_2$ we deduce that (v_n) is also Cauchy in $\mathcal{H}_2$ so there exists $\bar{h} \in \mathcal{H}_2$ such that $\|v_n - \bar{h}\|_2 \to 0$. Now observe that

$$h(x) = \lim_{n\to\infty} v_n(x) = \bar{h}(x),\ \ \forall x \in X.$$

Hence $h = \bar{h} \in \mathcal{H}_2$. □

Remark B.5.7. There is a more elementary proof the existence of a RKHS with a given reproducing kernel K. More precisely denote by V the subspace of $C(X)$ spanned by the functions $y \mapsto K_x(y)$. Given two functions $u, v \in V$

$$u = \sum_i u_i K_{x_i},\ \ v = \sum_j v_j K_{y_j},\ \ u_i, v_j \in \mathbb{R},$$

we define

$$\langle u, v\rangle = \sum_{i,j} u_i v_j K(x_i, y_j).$$

One can show that $\langle -,-\rangle$ is independent of the decompositions of u and v as linear combinations of functions K_x and it is *positive* definite and thus defines an inner product on V. We denote by $\mathcal{H}$ its completion. We can identify each $h \in \mathcal{H}$ with a function on X by setting $h(x) = \langle h, K_x\rangle$, $\forall x \in X$. □

For every $K \in \mathcal{K}(X)$ we denote by $\mathcal{H}_K$ the unique RKHS with reproducing kernel K. We denote by $(-,-)_K$ the inner product on $\mathcal{H}_K$.

We can now describe a few simple examples.

Example B.5.8. Suppose that X is a finite set $X = \{x_1, \dots, x_n\}$ and $K \in \mathcal{K}^+(X)$. We set $G = \mathbb{G}_K(x_1, \dots, x_n)$. Fix jointly Gaussian centered random variables $Z_1, \dots, Z_n$, with covariance matrix G and set

$$\mathcal{Z} = \operatorname{span}\big(Z_1, \dots, Z_n).$$

Let $(\delta_{x_i})_{1 \leq i \leq n}$ denote the canonical basis of $\mathbb{R}^X$,

$$\delta_{x_i}(x_j) = \delta_{ij}, \ \ \forall i, j.$$

Then the feature map is given by

$$x_i \mapsto K_{x_i} = \sum_j K(x_i, x_j)\delta_{x_j} \in \mathbb{R}^X.$$

Since $\langle K_{x_i}, K_{x_j} \rangle_{\mathcal{H}} = K(x_i, x_j)$ we deduce that in the basis (δ_{x_i}) the inner product $\langle -,- \rangle_{\mathcal{H}}$ is represented by the matrix by the matrix G^{-1}. □

Example B.5.9. Suppose that V is a Hilbert space with inner product $(-,-)$. Then the function

$$K : V \times V \to \mathbb{R}, \ \ K(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$$

is a kernel. The associated RKHS is the topological dual $V^* \subset \mathbb{R}^V$ equipped with the dual metric. The feature map is the Riesz representation isomorphism $V \to V^*$. □

Proposition B.5.10. *Suppose that $\mathcal{H}$ is a RKHS over X with reproducing kernel K. If $(e_i)_{i \in I}$ is a complete orthonormal system of $\mathcal{H}$, then*

$$K(x, y) = \sum_{i \in I} e_i(x) e_i(y).$$

Proof. We have

$$K_x = \sum_i (K_x, e_i)_{\mathcal{H}} e_i = \sum_i e_i(x) e_i$$

$$K(x, y) = (K_x, K_y)_{\mathcal{H}} = \sum_{i,j} e_i(x) e_j(y)(e_i, e_j) = \sum_i e_i(x) e_i(y).$$

□

Definition B.5.11. Let $\mathcal{H}$ be an RKHS over X with reproducing kernel K. A collection of functions $(f_i)_{i \in I}$ in $\mathcal{H}$ is called a *Parseval frame* if

$$\|h\|^2 = \sum_i \big| (h, f_i)_{\mathcal{H}} \big|^2, \ \ \forall f \in \mathcal{H}.$$

□

Clearly, a complete orthonormal collection of $\mathcal{H}$ is a Parseval frame. Parseval frames enjoy many of the properties of orthonormal bases. However a Parseval frame could have linearly dependent functions. We have the following useful characterization of Parseval frames.

Theorem B.5.12. *Let $\mathcal{H}$ be an RKHS over the topological space X with reproducing kernel K. A collection of continuous functions $(f_i)_{i \in I}$ on X is a Parseval frame of $\mathcal{H}_K$ if and only if*

$$K(x, y) = \sum_{i \in I} f_i(x) f_i(y),$$

where the above sum converges pointwisely. □

For a proof we refer to [**129**, Thm. 2.10, Remark 2.11]. Let us emphasize that above, *the functions f_i are not a priori known to belong to* $\mathcal{H}$! However, the above result implies that they span $\mathcal{H}_K$ and in fact

$$h = \sum_i (h, f_i)_{\mathcal{H}} f_i, \quad \forall h \in \mathcal{H}.$$

For more details we refer to [**129**, Sec. 2.1].

Appendix C

Probability

C.1. Gaussian random symmetric matrices

We denote by $\mathcal{S}_m$ the space of real symmetric $m \times m$ matrices. This is a Euclidean space with respect to the inner product $(A, B) := \operatorname{tr}(AB)$. This inner product is invariant with respect to the action of the orthogonal group $O(m)$ on $\mathcal{S}_m$.

We define

$$\ell_{ij}, \omega_{ij} : \mathcal{S}_m \to \mathbb{R}, \;\; \ell_{ij}(A) = a_{ij}, \;\; \omega_{ij}(A) := \begin{cases} a_{ij}, & i = j, \\ \sqrt{2}a_{ij}, & i < j. \end{cases}$$

The collection $(\omega_{ij})_{i \leq j}$ defines linear coordinates on $\mathcal{S}_m$ that are orthonormal with respect to the above inner product on $\mathcal{S}_m$. The volume density induced by this metric is

$$\operatorname{vol}\big[\, dA \,\big] := \prod_{i \leq j} d\omega_{ij} = 2^{\frac{1}{2}\binom{m}{2}} \prod_{i \leq j} d\ell_{ij}.$$

The space of $O(m)$-invariant homogeneous quadratic polynomials $q : \mathcal{S}_m \to \mathbb{R}$ is spanned by

$$q_1(A) := (\operatorname{tr} A)^2 \text{ and } q_2(A) := \operatorname{tr} A^2.$$

An $O(m)$-invariant homogeneous quadratic polynomial

$$q(A) = c_2 q_2(A) + c_1 q_1(A)$$

is nonnegative iff the quadratic form

$$F_q : \mathbb{R}^m \to \mathbb{R}, \;\; F_q\big(\lambda_1, \dots, \lambda_m\big) = c_2 \sum_k \lambda_k^2 + c_1 \Big(\sum_k \lambda_k \Big)^2$$

is nonnegative. This quadratic form is represented by the matrix

$$c_2 \mathbb{1}_m + c_1 S, \;\; s_{ij} = 1, \;\; \forall i, j.$$

Note that S has rank 1 and has only one nonzero eigenvalue m which is simple. We deduce that

$$c_2 q_2(A) + c_1 q_1(A) \geq 0 \Longleftrightarrow c_2 \geq 0, c_1 \geq -\frac{1}{m} c_2.$$

Note that

$$c_2 q_2(A) + c_1 q_1(A) = (c_2 + c_1) \sum_j \omega_{jj}^2 + 2c_1 \sum_{i<j} \omega_{ii}\xi_{jj} + c_2 \sum_{i<j} \omega_{ij}^2. \tag{C.1.1}$$

Throughout the book we encountered a 2-parameter family of Gaussian probability measures on $\mathcal{S}_m$. More precisely for any real numbers u, v such that

$$v > 0, mu + 2v > 0, \tag{C.1.2}$$

we denote by $\mathcal{S}_m^{u,v}$ the space $\mathcal{S}_m$ equipped with the centered Gaussian measure $\boldsymbol{\Gamma}_{u,v}\big[\, dA \,\big]$ uniquely determined by the covariance equalities

$$\mathbb{E}\big[\, \ell_{ij}(A)\ell_{k\ell}(A) \,\big] = u\delta_{ij}\delta_{k\ell} + v(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}), \;\; \forall 1 \leq i, j, \; k, \ell \leq m. \tag{C.1.3}$$

In particular we have

$$\mathbb{E}\big[\, \ell_{ii}^2 \,\big] = u + 2v, \;\; \mathbb{E}\big[\, \ell_{ii}\ell_{jj} \,\big] = u, \quad \mathbb{E}\big[\, \ell_{ij}^2 \,\big] = v, \;\; \forall 1 \leq i \neq j \leq m, \tag{C.1.4}$$

while all other covariances are trivial. The ensemble $\mathcal{S}_m^{0,v}$ is a rescaled version of the Gaussian Orthogonal Ensemble (GOE) and we will refer to it as GOE_m^v.

Comparing (C.1.1) with (C.1.4) we deduce that the covariance form of $\boldsymbol{\Gamma}_{u,v}$ corresponds to the $O(m)$-invariant quadratic form $c_2 q_2(A) + c_1 q_1(A)$, where

$$c_2 = 2v, \;\; c_1 = u.$$

The inequalities (C.1.2) guarantee that the covariance form is positive definite so that $\boldsymbol{\Gamma}_{u,v}$ is nondegenerate.

For $u > 0$ the ensemble $\mathcal{S}_m^{u,v}$ can be given an alternate description. More precisely a random $A \in \mathcal{S}_m^{u,v}$ can be described as a sum

$$A = B + \; X\mathbb{1}_m, \;\; B \in \mathrm{GOE}_m^v, \;\; X \in \boldsymbol{N}(0, u), \quad B \text{ and } X \text{ independent}.$$

We write this

$$\mathcal{S}_m^{u,v} = \mathrm{GOE}_m^v \,\hat{+} N(0, u)\mathbb{1}_m, \tag{C.1.5}$$

where $\hat{+}$ indicates a sum of *independent* variables.

The probability density $d\boldsymbol{\Gamma}_{u,v}$ has the explicit description

$$\boldsymbol{\Gamma}_{u,v}\big[\, dA \,\big] = \frac{1}{(2\pi)^{\frac{m(m+1)}{4}} \sqrt{D(u,v)}} e^{-\frac{1}{4v} \operatorname{tr} A^2 - \frac{u'}{2} (\operatorname{tr} A)^2} \operatorname{vol}\big[\, dA \,\big],$$

where

$$D(u, v) = (2v)^{(m-1) + \binom{m}{2}} \big(mu + 2v \,\big),$$

and

$$u' = \frac{1}{m} \left(\frac{1}{mu + 2v} - \frac{1}{2v} \right) = -\frac{u}{2v(mu + 2v)}.$$

In the special case GOE_m^v we have $u = u' = 0$ and

$$\boldsymbol{\Gamma}_{0,v}\big[\, dA \,\big] = \frac{1}{(4\pi v)^{\frac{m(m+1)}{4}}} e^{-\frac{1}{4v} \operatorname{tr} A^2} \operatorname{vol}\big[\, dA \,\big]. \tag{C.1.6}$$

Note that $\mathrm{GOE}_m^{1/2}$ corresponds to the Gaussian measure on $\mathbf{Sym}(\mathbb{R}^m)$ canonically associated to the inner product $(A, B) = \operatorname{tr}(AB)$.

We have a *Weyl integration formula* [5] which states that if $f : \mathcal{S}_m \to \mathbb{R}$ is a measurable function which is invariant under conjugation, then the value $f(A)$ at $A \in \mathcal{S}_m$ depends only on the eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_n(A)$ of A and we have

$$\mathbb{E}_{\mathrm{GOE}_m^v}\big[\, f(X)\,\big] = \frac{1}{\boldsymbol{Z}_m(v)} \int_{\mathbb{R}^m} f(\lambda_1, \dots, \lambda_m) \underbrace{\left(\prod_{1 \leq i < j \leq m} |\lambda_i - \lambda_j| \right) \prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}}_{=:Q_{m,v}(\lambda)} |d\lambda_1 \cdots d\lambda_m|, \tag{C.1.7}$$

where the normalization constant $\boldsymbol{Z}_m(v)$ is defined by

$$\boldsymbol{Z}_m(v) = \int_{\mathbb{R}^m} \prod_{1 \leq i < j \leq m} |\lambda_i - \lambda_j| \prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}} |d\lambda_1 \cdots d\lambda_m|$$

$$= (2v)^{\frac{m(m+1)}{4}} \times \underbrace{\int_{\mathbb{R}^m} \prod_{1 \leq i < j \leq m} |\lambda_i - \lambda_j| \prod_{i=1}^m e^{-\frac{\lambda_i^2}{2}} |d\lambda_1 \cdots d\lambda_m|}_{=:\boldsymbol{Z}_m}.$$

The integral $\boldsymbol{Z}_m$ is usually referred to as *Mehta's integral*. Its value was first determined in 1960 by M. L. Mehta, [98]. Later Mehta observed that this integral was known earlier to N. G. de Brujin [28]. It was subsequently observed that Mehta's integral is a limit of the *Selberg integrals*, [5, Eq. (2.5.11)], [61, Sec. 4.7.1]. We have

$$\boldsymbol{Z}_m = (2\pi)^{\frac{m}{2}} \prod_{j=0}^{m-1} \frac{\Gamma(\frac{j+3}{2})}{\Gamma(3/2)} = 2^{\frac{3m}{2}} \prod_{j=0}^{m-1} \Gamma\Big(\frac{j+3}{2} \Big). \tag{C.1.8}$$

In Section 2.3.4 we describe a probabilistic proof of this equality.

For any positive integer n we define the *normalized* 1-point correlation function $\rho_{n,v}(x)$ of GOE_n^v to be

$$\rho_{n,v}(x) = \frac{1}{\boldsymbol{Z}_n(v)} \int_{\mathbb{R}^{n-1}} Q_{n,v}(x, \lambda_2, \dots, \lambda_n) d\lambda_1 \cdots d\lambda_n.$$

For any Borel measurable function $f : \mathbb{R} \to \mathbb{R}$ we have [43, §4.4]

$$\frac{1}{n}\mathbb{E}_{\mathrm{GOE}_n^v}\big[\, \operatorname{tr} f(X)\,\big] = \int_{\mathbb{R}} f(\lambda)\rho_{n,v}(\lambda) d\lambda. \tag{C.1.9}$$

The equality (C.1.9) characterizes $\rho_{n,v}$. For example, if $f(x)$ is the indicator set of a Borel subset $B \subset \mathbb{R}$, then $\operatorname{tr} \boldsymbol{I}_B(X)$ the number of eigenvalues of X located in B so

$$\int_B \rho_{n,v}(\lambda) d\lambda$$

is the expected fraction of eigenvalues in B of a random matrix X in the ensemble GOE_n^v.

Let us observe that for any constant $c > 0$, if

$$A \in \mathrm{GOE}_n^v \Longleftrightarrow cA \in \mathrm{GOE}_n^{c^2 v}.$$

Hence for any Borel set $B \subset \mathbb{R}$ we have

$$\int_{cB} \rho_{n,c^2v}(x) dx = \int_B \rho_{n,v}(y) dy.$$

We conclude that

$$c\rho_{n,c^2v}(cy) = \rho_{n,v}(y), \quad \forall n, c, y. \tag{C.1.10}$$

We want to draw attention to a confusing situation in the existing literature on the subject. Some authors, such as M. L. Mehta [99], define the 1-point correlation function $R_n(x)$ by the equality

$$\mathbb{E}_{\mathrm{GOE}_n^{1/2}}\big[\,\operatorname{tr} f(X)\,\big] = \int_{\mathbb{R}} f(\lambda) R_n(\lambda) d\lambda.$$

so that

$$R_n(x) = n\rho_{n,1/2}(x). \tag{C.1.11}$$

From the equality

$$n^{1/2}\rho_{n,1/2}\big(\,n^{1/2}x\,\big) = \rho_{n,1/2n}(x)$$

we deduce

$$\rho_{n,1/2n}(x) = \frac{1}{\sqrt{n}} R_n\big(\,n^{1/2}x\,\big).$$

The expected value of the absolute value of the determinant of of a random $A \in \mathrm{GOE}_m^v$ can be expressed neatly in terms of the correlation function $\rho_{m+1,v}$. More precisely, we have the following result first observed by Y.V. Fyodorov [63] in a context related to ours.

Lemma C.1.1. *Suppose $v > 0$. Then for any $c \in \mathbb{R}$ we have*

$$\mathbb{E}_{\mathrm{GOE}_m^v}\big[\,|\det(A - c\mathbb{1}_m)|\,\big] = 2^{\frac{3}{2}}(2v)^{\frac{m+1}{2}}\Gamma\left(\frac{m+3}{2}\right) e^{\frac{c^2}{4v}}\rho_{m+1,v}(c). \tag{C.1.12}$$

Proof. Using Weyl's integration formula we deduce

$$\mathbb{E}_{\mathrm{GOE}_m^v}\big[\,|\det(A - c\mathbb{1}_m)|\,\big] = \frac{1}{\boldsymbol{Z}_m(v)}\int_{\mathbb{R}^m}\prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}|c-\lambda_i|\prod_{i\leq j}|\lambda_i - \lambda_j| d\lambda_1\cdots d\lambda_m$$

$$= \frac{e^{\frac{c^2}{4v}}}{\boldsymbol{Z}_m(v)}\int_{\mathbb{R}^m} e^{-\frac{c^2}{4v}}\prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}|c-\lambda_i|\prod_{i\leq j}|\lambda_i - \lambda_j| d\lambda_1\cdots d\lambda_m$$

$$= \frac{e^{\frac{c^2}{4v}}\boldsymbol{Z}_{m+1}(v)}{\boldsymbol{Z}_m(v)}\frac{1}{\boldsymbol{Z}_{m+1}(v)}\int_{\mathbb{R}^m} Q_{m+1,v}(c,\lambda_1,\dots,\lambda_m) d\lambda_1\cdots d\lambda_m$$

$$= \frac{e^{\frac{c^2}{4v}}\boldsymbol{Z}_{m+1}(v)}{\boldsymbol{Z}_m(v)}\rho_{m+1,v}(c) = \frac{e^{\frac{c^2}{4v}}(2v)^{\frac{m+1}{2}}\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\rho_{m+1,v}(c)$$

$$= e^{\frac{c^2}{4v}}(2v)^{\frac{m+1}{2}}\frac{(2\pi)^{1/2}\Gamma(\frac{m+3}{2})}{\Gamma(3/2)}\rho_{m+1,v}(c) =$$

$$= (2v)^{\frac{m+1}{2}}e^{\frac{c^2}{4v}}2^{3/2}\Gamma\Big(\frac{m+3}{2}\Big)\rho_{m+1,v}(c).$$

□

The above result was generalized in [**10**, Lemma 3.2.3] or [**11**, Lemma 3.3]. To state this generalization we need to recall some terminology. If $\mu, \nu \in \mathrm{Meas}(\mathbb{R})$ are two finite measures, then we define their convolution to be the finite measure $\mu_0 * \mu_1 \in \mathrm{Meas}(\mathbb{R})$ defined by

$$\mu * \nu\big[\, B\,\big] = \int_{\mathbb{R}} \mu\big[\, B - y\,\big]\nu\big[\, dy\,\big],\ \ \forall B \in \mathcal{B}_{\mathbb{R}}. \tag{C.1.13}$$

If μ is absolutely continuous with respect to the Lebesgue measure, $\mu\big[\, dx\,\big] = \rho_\mu(x)\boldsymbol{\lambda}\big[\, dx\,\big]$, then $\mu * \nu$ is also absolutely continuous with respect to the Lebesgue measure and

$$\mu * \nu\big[\, dx\,\big] = \rho_\mu * \nu(x)\boldsymbol{\lambda}\big[\, dx\,\big],\ \ \rho_\mu * \nu(x) = \int_{\mathbb{R}} \rho(x-y)\nu\big[\, dy\,\big].$$

Lemma C.1.2. *Let $u > 0$. Then*

$$\begin{aligned} \mathbb{E}_{\mathcal{S}_m^{u,v}}\big[\,|\det(A - c\mathbb{1}_m)|\,\big] &= 2^{\frac{3}{2}}(2v)^{\frac{m+1}{2}}\Gamma\left(\frac{m+3}{2}\right)\frac{1}{\sqrt{2\pi u}}\int_{\mathbb{R}} \rho_{m+1,v}(c-x)e^{\frac{(c-x)^2}{4v} - \frac{x^2}{2u}}dx. \\ &= 2^{\frac{3}{2}}(2v)^{\frac{m+1}{2}}\Gamma\left(\frac{m+3}{2}\right)\theta^+_{m+1,v} * \boldsymbol{\gamma}_u(c), \end{aligned} \tag{C.1.14}$$

where

$$\theta^+_{m+1,v}(x) = \rho_{m+1,v}(x)e^{\frac{x^2}{4v}}.$$

Let us observe that (C.1.12) can be obtained from (C.1.14) by letting $u \searrow 0$.

Proof. Recall the equality (C.1.5) $\mathcal{S}_m^{u,v} = \mathrm{GOE}_m^v \hat{+} \boldsymbol{N}(0,u)\mathbb{1}_m$. We deduce that

$$\begin{aligned} \mathbb{E}_{\mathcal{S}_m^{u,v}}\big[\,|\det(A - c\mathbb{1}_m)|\,\big] &= \mathbb{E}\big[\,\det(B + (X-c)\mathbb{1})|\,\big] \\ = \frac{1}{\sqrt{2\pi u}}&\int_{\mathbb{R}} \mathbb{E}_{\mathrm{GOE}_m^v}\big[\,|\det(B - (c-X)\mathbb{1}_m)|\,\|\ X = x\,\big]e^{-\frac{x^2}{2u}}dx \\ &= \frac{1}{\sqrt{2\pi u}}\int_{\mathbb{R}} \mathbb{E}_{\mathrm{GOE}_m^v}\big[\,|\det(B - (c-x)\mathbb{1}_m)|\,\big]e^{-\frac{x^2}{2u}}dx \\ = 2^{\frac{3}{2}}(2v)^{\frac{m+1}{2}}&\Gamma\left(\frac{m+3}{2}\right)\frac{1}{\sqrt{2\pi u}}\int_{\mathbb{R}} \rho_{m+1,v}(c-x)e^{\frac{(c-x)^2}{4v} - \frac{x^2}{2u}}dx. \end{aligned}$$

□

The behavior $\rho_{n,v}$ as $n \to \infty$ is controlled by the following theorem.

Theorem C.1.3 (Wigner's semicircle law)**.** *For any $v > 0$ the sequence of measures on $\mathbb{R}$*

$$\rho_{n,vn^{-1}}(x)dx = n^{\frac{1}{2}}\rho_{n,v}(n^{\frac{1}{2}}x)dx$$

converges weakly as $n \to \infty$ to the semicircle distribution

$$\rho_{\infty,v}(x)|dx| = \boldsymbol{I}_{\{|x| \le 2\sqrt{v}\}}\frac{1}{2\pi v}\sqrt{4v - x^2}|dx|.$$

□

For a proof we refer to [**5**, Chap. 2]. For our purposes we need a better understanding some refinements of Wigner's semicircle law. The following results were developed Set

$$\bar{\rho}_n(x) := \rho_{n,1/2n}(x) = n^{-1/2}R_n\big(\, n^{1/2}x\,\big),\ \ \rho(x) = \rho_{\infty,1/2}$$

where $R_n(x)$ is the 1-point correlation function $R_n(x)$ defined in (C.1.11).

Proposition C.1.4. *Denote by* $\boldsymbol{\gamma}_{vn^{-1}}$ *the centered Gaussian measure on* $\mathbb{R}$ *with variance* vn^{-1}*. Then*

$$\lim_{n\to\infty}\int_{\mathbb{R}}\bar{\rho}_n(s)\boldsymbol{\gamma}_{vn^{-1}}\big[\,ds\,\big]=\rho(0)=\frac{\sqrt{2}}{\pi}. \tag{C.1.15}$$

Proof. We follow the approach in [113]. The function $R_n(x)$ can be described explicitly in terms of Hermite[1] polynomials, [99, Eq. (7.2.32) and §A.9],

$$R_n(x)=\underbrace{\sum_{k=0}^{n-1}\psi_k(x)^2}_{=:\boldsymbol{k}_n(x)}+\underbrace{\left(\frac{n}{2}\right)^{\frac{1}{2}}\psi_{n-1}(x)\int_{\mathbb{R}}\varepsilon(x-t)\psi_n(t)dt+\alpha_n(x),}_{=:\boldsymbol{\ell}_n(x)} \tag{C.1.16}$$

where

$$\psi_n(x)=\frac{1}{(2^n n!\sqrt{\pi})^{\frac{1}{2}}}e^{-\frac{x^2}{2}}H_n(x),\quad H_n(x)=(-1)^n e^{x^2}\frac{d^n}{dx^n}(e^{-x^2}),$$

$$\alpha_n(x)=\begin{cases}0, & n\in 2\mathbb{Z},\\ \frac{\psi_{n-1}(x)}{\int_{\mathbb{R}}\psi_{n-1}(x)dx}, & n\in 2\mathbb{Z}+1,\end{cases}$$

and

$$\varepsilon(x)=\begin{cases}\frac{1}{2}, & x>0\\ 0, & x=0,\\ -\frac{1}{2}, & x<0.\end{cases}$$

From the Christoffel-Darboux formula [149, Eq. (5.5.9)] we deduce

$$\pi^{\frac{1}{2}}e^{x^2}\sum_{k=0}^{n-1}\psi_k(x)^2=\sum_{k=1}^{n-1}\frac{1}{2^k k!}H_k(x)^2=\frac{1}{2^n(n-1)!}\big(\,H_n'(x)H_{n-1}(x)-H_n(x)H_n'(x)\,\big)$$

Using the recurrence formula $H_n'=2xH_n-H_{n+1}$ we deduce

$$H_n'(x)H_{n-1}(x)-H_n(x)H_n'(x)=H_n^2(x)-H_{n-1}(x)H_{n+1}(x)$$

and

$$\boldsymbol{k}_n(x)=\frac{e^{-x^2}}{2^n(n-1)!\pi^{\frac{1}{2}}}\Big(H_n^2(x)-H_{n-1}(x)H_{n+1}(x)\Big).$$

We set

$$\bar{\boldsymbol{k}}_n(x):=\frac{\boldsymbol{k}_n\big(\sqrt{n}x\big)}{\sqrt{n}}=\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\psi_k\big(\,n^{1/2}x\,\big)^2$$

$$\bar{\boldsymbol{\ell}}_n(x):=\frac{\boldsymbol{\ell}_n\big(\sqrt{n}x\big)}{\sqrt{n}}=\frac{1}{\sqrt{2}}\psi_{n-1}\big(\,n^{1/2}x\,\big)\int_{\mathbb{R}}\varepsilon\big(\,n^{1/2}x-t\,\big)\psi_n(t)dt+\alpha_n\big(\,n^{1/2}x\,\big),$$

$$\bar{R}_n(x)=\frac{1}{\sqrt{n}}R_n(\sqrt{n}x)=\bar{\rho}_n(x)$$

so that

$$\bar{\rho}_n(x)=\bar{\boldsymbol{k}}_n(x)+\bar{\boldsymbol{\ell}}_n(x).$$

Lemma C.1.5.

$$\lim_{n\to\infty}\sup_{x\in\mathbb{R}}|\bar{\boldsymbol{\ell}}_n(x)|=0. \tag{C.1.17}$$

[1] In [99] the author uses a different convention for Hermite polynomials than the one we use in this book.

Proof. Using the generating series [**149**, Eq. (5.5.7)]

$$\sum_{n=0}^{\infty} H_n(x)\frac{T^n}{n!} = e^{2Tx-T^2}$$

we deduce that

$$\sum_{n=0}^{\infty}\left(\int_{\mathbb{R}} e^{-\frac{x^2}{2}} H_n(x)dx\right)\frac{T^n}{n!} = e^{T^2}\int_{\mathbb{R}} e^{-\frac{(x-2T)^2}{2}}dx = \sqrt{2\pi}e^{T^2},$$

so that

$$\frac{1}{(2n)!}\int_{\mathbb{R}} e^{-\frac{x^2}{2}} H_{2n}(x)dx = \frac{\sqrt{2\pi}}{n!} \text{ and } \int_{\mathbb{R}}\psi_{2n}(x)dx = \frac{\sqrt{2(2n)!}}{2^n n!\pi^{\frac{1}{4}}} \sim \text{const}\cdot n^{\frac{1}{4}} \text{ as } n\to\infty.$$

Using [**43**, Thm. 6.55] or [**149**, Thm. 8.91.3] we deduce that

$$\sup_{x\in\mathbb{R}}|\psi_n(x)| = O(n^{-\frac{1}{12}})$$

and thus

$$\sup_{x\in\mathbb{R}}|\alpha_n(x)| = O(n^{-\frac{1}{12}-\frac{1}{4}}) = O(n^{-\frac{1}{3}}) \text{ as } n\to\infty.$$

We set

$$F_n(x) = \int_{\mathbb{R}}\varepsilon(x-t)\psi_n(t)dt.$$

Using [**43**, Thm. 6.55 + Eq. (6.26)] we deduce $\sup_{x\in\mathbb{R}}|F_n(x)| = O\big(n^{-\frac{1}{12}}\big)$. This proves (C.1.17). □

Since $\boldsymbol{\gamma}_{vn^{-1}}$ converges weakly to the Dirac measure δ_0 we deduce from the above lemma and the uniform boundedness principle

$$\int_{\mathbb{R}}\big(\bar{\rho}_n(s)-\rho(s)\big)\boldsymbol{\gamma}_{vn^{-1}}\big[ds\big] = \int_{\mathbb{R}}\big(\bar{\boldsymbol{k}}_n(s)-\rho(s)\big)\boldsymbol{\gamma}_{vn^{-1}}\big[ds\big] + O\big(n^{-\frac{1}{12}}\big) \text{ as } n\to\infty.$$

Lemma C.1.6.

$$\lim_{n\to\infty}\int_{\mathbb{R}}\big(\bar{\boldsymbol{k}}_n(s)-\rho(s)\big)\boldsymbol{\gamma}_{vn^{-1}}\big[ds\big] = 0.$$

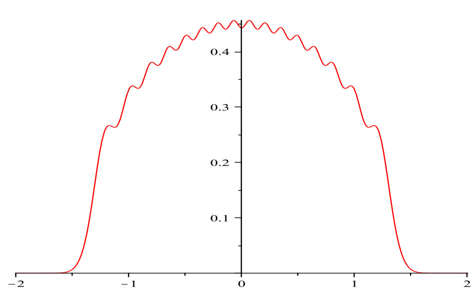

Figure C.1. *The graph of* $\bar{\boldsymbol{k}}_{16}(x)$, $|x|\le 2$.

Proof. Denote by $w_n(s)$ the density of the Gaussian measure $\boldsymbol{\gamma}_{vn^{-1}}\big[\,ds\,\big]$,

$$w_n(s) = \frac{n^{1/2}}{\sqrt{2\pi v}} e^{-\frac{vs^2}{2v}}.$$

Fix $c \in (0, \sqrt{2})$ so that the interval $(-c, c)$ lies inside the oscillatory regime of $H_n(\sqrt{n}t)$. We have

$$\begin{aligned} &\int_{\mathbb{R}} \big(\,\bar{\boldsymbol{k}}_n(s) - \rho(s)\,\big) w_n(s)ds \\ &= \int_{|s|\leq c} \big(\,\bar{\boldsymbol{k}}_n(s) - \rho(s)\,\big) w_n(s)ds + \int_{|s|>c} \big(\,\bar{\boldsymbol{k}}_n(s) - \rho(s)\,\big) w_n(s)ds \\ &\leq \sup_{|s|\leq c} |\bar{\boldsymbol{k}}_n(s) - \rho(s)| + \sup_{|s|>c} |\big(\,\bar{\boldsymbol{k}}_n(s) - \rho(s)\,\big)| \int_{|s|>c} w_n(s)ds. \end{aligned}$$

Using the Plancherel-Rotach formulæ ([**43**, Eq. (6.126)], [**133**], [**149**, Thm. 8.22.9]) and arguing as in [**61**, §7.1.6] or [**64**, §6.1] we deduce that

$$\lim_{n\to\infty} \sup_{|s|\leq c} |\bar{\boldsymbol{k}}_n(s) - \rho(s)| = 0.$$

On the other hand

$$\lim_{n\to\infty} \int_{|s|>c} w_n(s)ds = 0,$$

and [**149**, Thm.8.91.3] implies that

$$\sup_{|s|>c} |\big(\,\bar{\boldsymbol{k}}_n(s) - \rho(s)\,\big)| = O(1) \;\; \text{as } n \to \infty.$$

□

Since $\boldsymbol{\gamma}_{vn^{-1}}\big[\,ds\,\big]$ converges to the Dirac measure δ_0 we deduce again from the uniform boundedness principle tha

$$\lim_{n\to\infty} \int_{\mathbb{R}} \rho(s) w_n(s) \boldsymbol{\gamma}_{vn^{-1}}\big[\,ds\,\big] = \rho(0) = \frac{\sqrt{2}}{\pi}.$$

□

C.2. Random measures

Denote by $\widehat{\mathrm{Meas}}\big(\mathbb{R}^m\big)$ the space locally finite of Borel measures μ on $\mathbb{R}^m$, i.e., $\mu\big[\,B\,\big] < \infty$ for any bounded Borel subset $B \subset \mathbb{R}^m$. Each $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$ defines a map

$$I_f : \widehat{\mathrm{Meas}}\big(\mathbb{R}^m\big) \to \mathbb{R}, \;\; \mu \mapsto I_f(\mu) = \mu\big[\,f\,\big] := \int_{\mathbb{R}^m} f(\boldsymbol{x}) \mu\big[\,dx\,\big].$$

The *vague topology* on $\widehat{\mathrm{Meas}}\big(\mathbb{R}^m\big)$ is the smallest topology such that all the functions I_f, $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$ are continuous. As shown in [**79**, Thm. 4.2], this topology is Polish, i.e., it is induced by a complete and separable metric. We denote by $(\mathcal{M}, d)$ this metric space. The convergence in this metric is called *vague convergence*.

A sequence of measure (μ_n) in $\mathcal{M}$ converges vaguely to $\mu \in \mathcal{M}$, and we indicate this as $\mu_n \xrightarrow{v} \mu$, if and only if

$$\mu_n\big[\,f\,\big] \to \mu\big[\,f\,\big], \;\; \forall f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m).$$

A locally finite random measure on $\mathbb{R}^m$ is a Borel measurable map

$$\mathfrak{M} : \big(\Omega, \mathcal{S}, \mathbb{P}\big) \to \mathcal{M}.$$

Its distribution is a Borel probability measure on $\mathcal{M}$, $\mathbb{P}_{\mathfrak{M}} \in \operatorname{Prob}(\mathcal{M})$.

Recall that a sequence of probability measures $\mu_n \in \operatorname{Prob}(\mathcal{M})$ is said to converge *weakly* to $\mu \in \operatorname{Prob}(\mathcal{M})$, and we indicate this $\mu_n \to \mu$ if

$$\lim_{n\to\infty} \int_{\mathcal{M}} F d\mu_n = \int_{\mathcal{M}} F d\mu,$$

for any bounded and continuous function $F : \mathcal{M} \to \mathbb{R}$.

A sequence of random measures $\mathfrak{M}_n$ is said to converge weakly to the random measure $\mathfrak{M}$ if the distributions $\mathbb{P}_{\mathfrak{M}_n}$ converge weakly in $\operatorname{Prob}(\mathfrak{X})$ to $\mathbb{P}_{\mathcal{M}}$. We we use the notation $\mathfrak{M}_n \to \mathfrak{M}$ to indicate this. We have the following result, [**41**, Prop.11.1.VIII], [**79**, Thm. 4.11].

A subset $Q \subset \mathbb{R}^m$ is called a *quasi-box* if it is a product of finite intervals

$$Q = I_1 \times \cdots \times I_m.$$

The intervals I_k need not be closed and could have length zero. Note that a quasi-box Q is a box if all the intervals I_k are closed and have nonzero lengths.

Theorem C.2.1. *Consider a sequence* $\big(\mathfrak{M}_n\big)_{n\in\mathbb{N}}$ *of random locally finite measures on* $\mathbb{R}^m$*. The following are equivalent.*

(i) *The sequence* $\mathfrak{M}_n$ *converges weakly to the random locally finite measure* $\mathfrak{M}$*.*

(ii) *For any* $f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$*, the random variables* $\mathfrak{M}_n\big[\, f\,\big]$ *converge in distribution to* $\mathfrak{M}\big[\, f\,\big]$*.*

(iii) *For any quasi-box* $Q \subset \mathbb{R}^m$ *the random variables* $\mathfrak{M}_n\big[\, Q\,\big]$ *converge in distribution to* $\mathfrak{M}\big[\, Q\,\big]$*.*

□

There are other modes of convergence of random measures corresponding to the various modes of convergence of random variables. Suppose that

$$\mathfrak{M}_n,\ \mathfrak{M} : \big(\Omega, \mathcal{S}, \mathbb{P}\big) \to \mathcal{M},\ \ n \in \mathbb{N}$$

are random locally finite measures. We say that $\mathfrak{M}_n$ *converges almost surely* to $\mathfrak{M}$, and we indicate this $\mathfrak{M}_n \overset{\text{a.s.}}{\to} \mathfrak{M}$, if there exists a $\mathbb{P}$-negligible set $\mathcal{N} \in \mathcal{S}$ such that

$$\mathfrak{M}_n(\omega) \overset{v}{\to} \mathfrak{M}(\omega),\ \ \forall \omega \in \Omega \setminus \mathcal{N},$$

i.e.,

$$\mathfrak{M}_n \xrightarrow{\text{a.s.}} \mathfrak{M} \Longleftrightarrow \mathfrak{M}_n\big[\, f\,\big] \xrightarrow{\text{a.s.}} \mathfrak{M}\big[\, f\,\big],\ \ \forall f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m).$$

The convergence $\mathfrak{M}_n \xrightarrow{L^p} \mathfrak{M}$ is defined in a similar fashion namely

$$\mathfrak{M}_n \xrightarrow{L^p} \mathfrak{M} \Longleftrightarrow \mathfrak{M}_n\big[\, f\,\big] \xrightarrow{L^p} \mathfrak{M}\big[\, f\,\big],\ \forall f \in C^0_{\mathrm{cpt}}(\mathbb{R}^m).$$

One can show (see [**79**, Lemma 4.8]) that

$$\mathfrak{M}_n \xrightarrow{\text{a.s.}} \mathfrak{M} \Longleftrightarrow \mathfrak{M}_n\big[\, Q\,\big] \xrightarrow{\text{a.s.}} \mathfrak{M}\big[\, Q\,\big],\ \ \text{for any quasi-box } Q \subset \mathbb{R}^m, \tag{C.2.1a}$$

$$\mathfrak{M}_n \xrightarrow{L^p} \mathfrak{M} \Longleftrightarrow \mathfrak{M}_n\big[\, Q\,\big] \xrightarrow{L^p} \mathfrak{M}\big[\, Q\,\big],\ \ \text{for any quasi-box } Q \subset \mathbb{R}^m. \tag{C.2.1b}$$

The action of $\mathbb{R}^m$ on itself by translations induces and action on $\mathcal{M} = \widehat{\mathrm{Meas}}\big(\mathbb{R}^m\big)$,

$$\mathcal{T} : \mathbb{R}^m \times \mathcal{M} \to \mathcal{M}, \;\; \mathbb{R}^m \times \mathcal{M} \ni (\boldsymbol{x}, \mu) \mapsto \mathcal{T}_{\boldsymbol{x}}\mu,$$

where $\mathcal{T}_x\mu\big[\,B\,\big] = \mu\big[\,B - \boldsymbol{x}\,\big]$, for any Borel subset $B \subset \mathbb{R}^m$. We denote by $\mathcal{I}$ the sigma-subalgebra of $\mathcal{B}_{\mathcal{M}}$ consisting of Borel subsets of $\mathcal{M}$ that are invariant with respect to the above action. A measure $\mathbb{P} \in \mathrm{Prob}\big(\mathcal{M}\big)$ is called *stationary* if it is invariant with respect to this action.

$$(\mathcal{T}_{\boldsymbol{x}})_{\#}\mathbb{P} = \mathbb{P}, \;\; \forall \boldsymbol{x} \in \mathbb{R}^m.$$

A random measure $\mathfrak{M}$ is called *stationary* if its distribution $\mathbb{P}_{\mathfrak{M}}$ is stationary.

Wiener's ergodic theorem [80, Thm. 25.4], [160] shows that if $\mathbb{P} \in \mathrm{Prob}\big(\mathcal{M}\big)$ is stationary, then for any $F \in L^1\big(\mathcal{M}, \mathbb{P}\big)$ and any compact convex set $C \subset \mathbb{R}^m$ containing the origin in the interior we have

$$\lim_{N\to\infty} \frac{1}{\mathrm{vol}\,\big[\,NC\,\big]} \int_{NC} \mathcal{T}^*_{\boldsymbol{x}} F \boldsymbol{\lambda}\big[\,d\boldsymbol{x}\,\big] = \mathbb{E}_{\mathbb{P}}\big[\,F \,\|\, \mathcal{I}\,\big], \tag{C.2.2}$$

$\mathbb{P}$-a.s. and L^1.

Let C_1 denote the unit cube in $\mathbb{R}^m$. If $\mathfrak{M}$ is a stationary random locally finite measure on $\mathbb{R}^m$ such that $\mathbb{E}\big[\,\mathfrak{M}[C_1]\,\big] < \infty$, we define its *asymptotic intensity* to be the random variable

$$\overline{\mathfrak{M}} := \mathbb{E}\big[\,\mathfrak{M}[C_1] \,\|\, \mathcal{I}\,\big]$$

The Wiener's ergodic theorem applied to the action of $\mathbb{Z}^m \subset \mathbb{R}^m$ on $\mathcal{M}$ implies (see [79, Th. 5.23] and [152, Thm. 6.4.1]) that for any compact convex subset $C \subset \mathbb{R}^m$ containing the origin in the interior we have

$$\frac{1}{N^m}\mathfrak{M}\big[\,NC\,\big] \to \overline{\mathcal{M}} \cdot \mathrm{vol}\,\big[\,C\,\big] \tag{C.2.3}$$

a.s. and L^1. Moreover, if $\mathfrak{M}[C_1] \in L^p$, then the convergence holds also in L^p.

A sequence $\varphi_N \in C^0_{\mathrm{cpt}}(\mathbb{R}^m)$, $N \in \mathbb{N}$ is called *asymptotically stationary* if

$$\varphi_N \geq 0, \;\; \int_{\mathbb{R}^m} \varphi_N(\boldsymbol{x}) d\boldsymbol{x} = 1, \;\; \forall N,$$

and

$$\lim_{N\to\infty} \int_{\mathbb{R}^m} \big|\,\varphi_N(\boldsymbol{x}) - \varphi_N(\boldsymbol{x} - \boldsymbol{y})\,\big| d\boldsymbol{x} = 0, \;\; \forall \boldsymbol{y} \in \mathbb{R}^m.$$

We have the following result, [79, Thm. 5.24] and [152, Thm. 6.4.1].

Theorem C.2.2. *If $\mathfrak{M}[C_1] \in L^p$, $p \in [1, \infty)$, and $(\varphi_N)_{N\in\mathbb{N}}$ is asymptotically stationary, then*

$$\mathfrak{M}\big[\,\varphi_N\,\big] \to \overline{\mathfrak{M}},$$

in L^p and a.s.. □

Bibliography

[1] R. Adler, R.J.E. Taylor: *Random Fields and Geometry*, Springer Monographs in Mathematics, Springer Verlag, 2007.

[2] M. Ancona, L. Gass, T. Letendre, M. Stecconi: *Zeros and Critical Points of Gaussian Fields: Cumulants Asymptotics and Limit Theorems*, arXiv: 2501.10226

[3] M. Ancona, T. Letendre: *Zeros of smooth stationary Gaussian processes*, Electron. J. Probab., **26**(2021), 1- 81, arXiv: 2007.03240

[4] M. Ancona, T. Letendre: *Multijet bundles and applications to the finiteness of moments for zeros of Gaussian fields*, arXiv.2307.10659

[5] G. W. Anderson, A. Guionnet, O. Zeitouni: *An Introduction to Random Matrices*, Cambridge University Press, 2010.

[6] M. Arcones: *Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors*, Ann. of Probability, **22**(1994), 2242-2274.

[7] D. Armentao, J.-M. Azaïs, J.-R. León: *On the general Kac-Rice formula for the measure of a level set*, arXiv: 2304.0742v1.

[8] V.I. Arnold: *Statistics and classification of topologies of periodic functions and trigonometric polynomials* Proc. Steklov Inst. Math. 2006, Dynamical Systems: Modeling, Optimization, and Control, suppl. 1, S13–S23.

[9] N. Aronszain: *Theory of reproducing kernels*, Trans. A.M.S. **68**(1950), 337-404.

[10] A. Auffinger: *Random matrices, complexity of spin glasses and heavy tailed processes*, 2011 NYU PhD Dissertation.

[11] A. Auffinger, G. Ben Aurous, J. Cerny: *Random matrices and complexity of spin glasses*, Comm. Pure Appl. Math., **66**(2013), 165-201. arXiv: 1003.1129v2

[12] J.-M. Azaïs, M. Wschebor: *Level Sets and Extrema of Random Processes*, John Wiley & Sons, 2009.

[13] J.-M. Azaïs, F. Dalmao, C. Delmas: *Multivariable CLT for critical points*, arXiv: 2401.09117.

[14] E. Azmoodeh, G. Peccati, X. Yang: *Malliavin-Stein method: a survey of some recent developments*, Mod. Stoch. Theory Appl., **8**(2021), 141-177.

[15] P. Baxendale: *Gaussian measures on function spaces*, Amer. J. Math., **98**(1976), 891-952.

[16] D. Beliaev, M. McAuley, S. Muirhead: *A central limit theorem for the number of excursion set components of Gaussian fields*, arXiv.2205.0985

[17] Yu. K. Belyaev: *Continuity and Hölder's conditions for sample functions of stationary Gaussian processes*, Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. II, pp. 23–33, Univ. California Press, Berkeley-Los Angeles, Calif., 1960.

[18] D. N. Bernshtein: *The number of roots of a system of equations*, Funct. Anal. and its Appl. , **9**(1975), no. 3, 1-4.

[19] Z. W. Birnbaum: *An inequality for Mills ratio*, Ann. Math. Statist., **2**(1942), 245-246.

[20] J. R. Blum, B. Eisenberg: *Conditions for metric transitivity for stationary Gaussian processes on groups*, Ann. Math. Studies, **43**(1972), 1737-1741.

[21] V. I. Bogachev: *Gaussian Measures*, Amer. Math. Soc., 1998.

[22] V. I. Bogacev: *Measure Theory, Vol.1*, Springer Verlag, 2007.

[23] V. I. Bogacev: *Measure Theory, Vol.2*, Springer Verlag, 2007.

[24] B.F. Bojanov, H. A. Hakopian, A. A. Sahakian: *Spline Functions and Multivariate Interpolations*, Kluwer, 1993.

[25] N. Bouleau, F. Hirsch: *Dirichlet Forms and Analysis on Wiener Space*, Walter de Gruyter, 1991.

[26] M. Brion, M. Vergne: *Residue formulæ, vector partition functions, and lattice points in rational polytopes*, J. Amer. Math. Soc., **10**(1997), 797-833.

[27] H. Brezis: *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer Verlag, 2011.

[28] N.G. de Brujin: *On some multiple integrals involving determinants*, J. Ind. Math. Soc., **19**(1955), 133-151.

[29] N.G. de Brujin: *Asymptotic Methods in Analysis*, Dover Publications, 1981.

[30] W. Bryc: *The Normal Distribution. Characterization with Applications*, Lect. Notes. in Statistics, vol. 100, Springer Verlag, 1995.

[31] Y. D. Burago, V.A. Zalgaller: *Geometric Inequalities*, Springer Verlag, 1988.

[32] R. H. Cameron, W. T. Martin: *Transformations of Wiener integrals under translations*, Ann. Math., **45**(1944), 386-396.

[33] S. Chatterjee: *A short survey of Stein's method*, 2014 ICM talk, arXiv: 1404.1392.

[34] L.H.Y. Chen, L. Goldstein, Q,-M, Shao: *Normal Approximations by Stein's Method* Springer Verlag, 2011.

[35] E. Çinlar: *Probability and Stochastics*, Grad. Texts in Math., vol. 261 Springer Verlag, 2011.

[36] D. L. Cohn: *Measure Theory*, 2nd Edition, Birkhäuser, 2013.

[37] H. Cramér, M. R. Leadbetter: *Stationary and Related Stochastic processes. Sample Function Properties and Their Applications*, John Wiley & Sons, 1967, Dover 2006.

[38] . S. D. Cutkowski: *Resolution of Singularities*, Grad. Stud. Math, vol. 63, Amer. Math. Soc., 2004.

[39] G. Da Prato: *An Introduction to Stochastic Analysis and Malliavin Calculus*, Edizione della Normale, 2014.

[40] D. J. Daley, D. Vere-Jones, *An Introduction to the Theory of Point Processes. Vol I: Elementary Theory and Methods*, 2nd Edition, Springer Verlag, 2003.

[41] D. J. Daley, D. Vere-Jones, *An Introduction to the Theory of Point Processes. Vol II: General Theory and Structure*, 2nd Edition, Springer Verlag, 2008.

[42] M. Das: *Real zeros of a random sum of orthogonal polynomials*, Proc. Amer. Math. Soc., **27**(1971), 147–153.

[43] P. Deift, D. Gioev: *Random Matrix Theory: Invariant Ensembles and Universality*, Courant Lecture Notes, vol. 18, Amer. Math. Soc., 2009.

[44] Y. Do, O. Nguyen, V. Vu: *Roots of random polynomials with coefficients with polynomial growth*, Ann. Prob., **46**(2018), 2407-2494, arXiv: 1507.04994

[45] J. L. Doob: *Stochastic Processes*, John Wiley & Sons, 1953.

[46] B. K Driver: *Probability Tools with Examples*, Course notes, 2010.

[47] B. K Driver: *Heat kernel weighted L^2-spaces*, unpublished notes,

[48] R. M. Dudley: *Sizes of compact subsets of Hilbert space and continuity of Gaussian processes*, J. Funct. Anal. **1**(1967), 290-330.

[49] R. M. Dudley: *Real Analysis and Probability*, Cambridge University Press, 2002.

[50] J. J. Duistermaat, J. A. C. Kolk: *Distributions. Theory and Applications*, Birkhäuser, 2010.

[51] A. Edelman, E. Kostlan: *How many zeros of a random polynomial are real?*, Bull. A.M.S. (New Series), **32**(1995), No.1, 1-37.

[52] R. E. Edwards: *Functional Analysis. Theory and Applications*, Holt, Reinhart & Winston, 1965.

[53] K. D. Elworthy, Y. LeJan, X.-M. Lin: *On the Geometry of Diffusion Operators and Stochastic Flows*, Lect. Notes in Math., vol. 1720, Springer Verlag, 1999.

[54] D. D. Engel: *The multiple stochastic integral*, Mem. A.M.S., vol. 38, no. 265, Amer. Math. Soc., 1982.

[55] A. Estrade, J. R. León: *A central limit theorem for the Euler characteristic of a Gaussian excursion set*, Ann. of Probab., **44**(2016), 3849-3878. https://hal.archives-ouvertes.fr/hal-00943054v3

[56] I. Fáry: *Sur la courbure totale d'une courbe gauche faisant un noed*, Bull. Soc. Math. France, **77**(1949), 128-138.

[57] H. Federer: *Geometric Measure Theory*, Springer Verlag, 1996. (Reprint of the 1969 edition.)

[58] W. Feller: *An Introduction to Probability Theory and its Applications*, Volume 1, 3rd Edition, John Wiley & Sons, 1970.

[59] W. Feller: *An Introduction to Probability Theory and Its Applications*, vol.2, John Wiley& Sons, 1970.

[60] X. Fernique: *Regularité des trajectoires des fonctions aléatoires gaussiennes*, École d'Étè de Probabilités de Saint Flour IV-1974, Lect. Notes. in Math. vol. 480, p. 1-96, Springer Verlag, 1975.

[61] P. J. Forrester: *Log-Gases and Random Matrices*, London Math. Soc. Monographs, Princeton University Press, 2010.

[62] Q. Fu. L. I. Nicolaescu: *A functional law of large numbers concerning the distribution of critical points of random Fourier series*

[63] Y. V. Fyodorov: *Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices*, Phys. Rev. Lett, **92**(2004), 240601; Erratum: **93**(2004), 149901.

[64] Y. V. Fyodorov: *Introduction to random matrix theory: Gaussian unitary ensemble and beyond*, in the volume *Recent perspectives in random matrix theory and number theory*, London Math. Soc. Lecture Note Ser., 322, Cambridge Univ. Press, Cambridge, 2005.

[65] L. Gass: *Cumulants asymptotics of the zeros counting measure of real Gaussian processes*, Electron. J. Probab., **28**(2023), 1-45, arXiv: 2112.08247

[66] L. Gass, M. Stecconi: *The number of critical points of a Gaussian field: finiteness of moments*, Probab. Theory Relat. Fields, **190**(2024), 167-197, arxiv.org: 2305.17586v2.

[67] I. M. Gelfand, N. Ya. Vilenkin: *Generalized Functions. Volume 4. Applications of Harmonic Analysis*, Academic Press, 1964.

[68] I. I. Gikhman, A. V. Skhorohod: *Introduction to the Theory of Random Processes*, Dover, 1996.

[69] E. Giné, R. Nickl: *Mathematical Foundations of Infinite Dimensional Statistical Models*, Cambridge University Press, 2016.

[70] J. L. van Hemmen, J. Ando: *An inequality for trace ideals* Comm. Math. Phys., **76**(1980), 143-148.

[71] M. W. Hirsch: *Differential Topology*, Springer Verlag, 1976.

[72] W. Holsztyński, J. Mycielski, G. C. O'Brien, G. C., S. Świerczkowski: *Closed curves on spheres*, Amer. Math. Monthly, **115**(2008), 149-153.

[73] L. Hörmander: *On the spectral function of an elliptic operator*, Acta Math. **121**(1968), 193-218.

[74] L. Hörmander: *The Analysis of Linear Partial Differential Operators I.*, Springer Verlag, 2003.

[75] R. A. Horn, C. R. Johnson: *Matrix Analysis*, Cambridge University Press, 1990 Reprint.

[76] S. Janson: *Gaussian Hilbert Spaces*, Cambridge University Press, 1997.

[77] M. Kac: *On the average number of real roots of a random algebraic equation*, Bull. A.M.S., **49**(1943), 314-320.

[78] J. P. Kahane: *A century of interplay between Taylor series, Fourier series and Brownian motion*, Bull. London. Math. Soc. **29**(1997), 256-267.

[79] O. Kallenberg: *Random measures, Theory and Applications*, Springer Verlag 2017.

[80] O. Kallenberg: *Foundations of modern probabiliity*, 3rd Edition, Springer Verlag, 2021.

[81] T. Kato: *Perturbation Theory for Linear Operators*, Classics in Mathematics, Springer Verlag, 1995.

[82] P. Kergin: *A natural interpolation of C^k-functions*, J. of Approx. Th., **29**(1980), 278-293.

[83] A.G. Kouchnirenko: *Polyèdres de Newton et nombres de Milnor*, Invent. Math. **32**(1976), 1-31.

[84] H.-H. Kuo: *Introduction to Stochastic Integration*, Springer Verlag, 2006.

[85] S. Kwapień, W. A. Woyczyński: *Random Series and Stochastic Integrals: Single and Multiple*, Birkhäuser, 1992.

[86] N. N. Lebedev: *Special Functions and Their Applications*, Dover, 1972.

[87] M. Ledoux: *The Concentration of Measure Phenomenon*, Math. Surveys and Monographs, vol. 89, Amer. Math. Soc., 2001.

[88] M. Ledoux, M. Talagrand: *Probability in Banach Spaces*, Springer Verlag, 1991.

[89] A. Lerario, M. Stecconi: *Differential Topology of Gaussian Random Fields*, arXiv: 1902.03805

[90] C. Ley, G. Reinert, Y. Swan: *Stein method for comparison of univariate distributions*, Probability Surveys, **14**(2017), 1-52.

[91] R. Lyons: *Strong laws of large numbers for weakly correlated random variables*, Mich. Math. J., **35**(1988), 353-359.

[92] D. R MacIver *Journal of Obscure Results 1: Nedoma Pathology* `https://www.drmaciver.com/2006/04/journal-of-obscure-results-1-nedomas-pathology/`

[93] P. Major: *Multiple Wiener-Ito integrals*, Lect. Notes in Math., vol. 849, Springer Verlag, 1981.

[94] P. Malliavin: *Integration and Probability*, Springer Verlag, 1997

[95] P. Malliavin: *Stochastic Analysis*, Grundlehren. der math. Wissenschaften, vol. 313, Springer Verlag, 1997.

[96] P. Malliavin, A. Thalmaier: *Stochastic Calculus of Variations in Mathematical Finance*, Springer Verlag, 2006.

[97] H. P. McKean: *Stochastic Integrals*, Academic Press, 1969.

[98] M. L. Mehta: *On the statistical proerties of level-spacings in nuclear spectra*, Nucl. Phys.. **18**(1960), 395-419.

[99] M. L. Mehta: *Random Matrices*, 3rd Edition, Elsevier, 2004.

[100] T. Mercer: *Functions of positive and negative type and their connections with the theory of integral equations*, Trans. London Phil. Soc., **209**(1909), 415-446.

[101] C. A. Miccheli, P. Milman: *A formula for Kergin interpolation in* $\mathbb{R}^k$, J. of Approx. Th., **29**(1980), 294-296.

[102] L. M. Milne-Thompson: *The Calculus of Finite Differences*, Macmillan & Co. 1933.

[103] Milnor, J. W. (1950). On the total curvature of knots. *Ann. Math.*, 52(2): 248–237.

[104] F. Móricz: *Multiparameter strong laws of large numbers. I (Second order moment restrictions)*, Acta. Sci. Math., **40**(1978), 143-156.

[105] C. Müller: *Analysis of Spherical Symmetries in Euclidean Spaces*, Appl. Math. Sci. vol. 129, Springer Verlag, 1998.

[106] F. J. Murray, J. v. Neumann: *Rings of operators*, Ann. Math., **37**(1936), 116-229.

[107] F. Nazarov, M. Sodin: *Asymptotic laws for the spatial distribution and the number of components of zero sets of Gaussian random functions*, J. Math. Phys., Analysis, Geom., **12**(2016), 205-278. `https://core.ac.uk/download/pdf/159626247.pdf`

[108] L. I. Nicolaescu: *An Invitation to Morse Theory*, 2nd Edition, Springer Verlag, 2011.

[109] L. I. Nicolaescu: *The Co-area formula*, `https://www3.nd.edu/~lnicolae/Coarea.pdf`

[110] L. I. Nicolaescu: *Critical sets of random smooth functions on products of spheres*, arXiv: 1008.5085.

[111] L. I. Nicolaescu: *Complexity of random smooth functions on compact manifolds*, Indiana J. Math., **63**(2014), 1037-1065.

[112] L. I. Nicolaescu: *Random Morse functions and spectral geometry*, arXiv: 12.090639.

[113] L. I. Nicolaescu: *Critical sets of random smooth functions on compact manifolds.*, Asian J. Math., **19**(2015), 391-432. arXiv: 1101.5990,

[114] L. I. Nicolaescu: *A stochastic Gauss-Bonnet-Chern formula*, Probab. Theory Relat. Fields, **165**(2016), 235-265.

[115] L. I, Nicolaescu: *Critical points of multidimensional random Fourier series: variance estimates*, J. Math. Phys., **57**(2016), 083304.

[116] L. I. Nicolaescu: *Critical points of multidimensional random Fourier series: central limits*, Bernoulli, **24**(2018), 1128-1170, arXiv:1511.04965v2.

[117] L. I. Nicolaescu: *Lectures on the Geometry of Manifolds*, 3rd Edition, World Scientific, 2021.

[118] L. I. Nicolaescu: *A Graduate Course in Probability*, World Scientific, 2022. `https://www3.nd.edu/~lnicolae/Grad_Prob_web.pdf` .

[119] L. I. Nicolescu: *Counting zeros of random functions*, Amer.Math. Monthly, **130**(2023), 625-646.

[120] L. I. Nicolaescu: *A probabilistic computation of a Mehta integral*, arXiv: 2408.06203.

[121] L. I. Nicolaescu: *A law of large numbers concerning the distribution of critical points of isotropic Gaussian functions*, arXiv: 2408.14383.

[122] L. Nirenberg: *Topics in Nonlinear Functional Analysis*, Amer. Math. Soc., 2001.

[123] I. Nourdin, G. Peccati: *Stein's method on Wiener chaos*, Prob. Theor. Rel. Fields, **145**(2009), 75-118.

[124] I. Nourdin, G. Peccati: *Normal Approximations with Malliavin Calculus. From Stein's Method to Universality*, Cambridge Tracts in Math., vol.192, Cambridge University Press, 2012.

[125] I. Nourdin, G. Peccati, M. Podolskij: *Quantitative Breuer-Major theorems*, Stoch. Proc. and their Appl., **121**(2011), 793-812.

[126] D. Nualart: *The Malliavin Calculus and Related Topics*, 2nd Edition, Springer Verlag, 2006.

[127] D. Nualart, G. Peccati: *Central limit theorems for sequences of multiple stochastic integrals*, Ann. Probab. **33**(2005), 177-193.

[128] G. di Nunno, B. Øksendal, F. Proske: *Malliavin Calculus for Lèvt Processes with Applications to Finance*, Springer Verlag, 2009.

[129] V. I. Paulsen, M. Raghupathi: *An Introduction to the Theory of Reproducing Kernel Spaces*, Cambridge University Press, 2016.

[130] A. Pazy: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer Verlag, 1983.

[131] G. Peccati, M. S. Taqqu: *Wiener Chaos: Moments, Cumulants and Diagrams. A Survey with Computer Implementation*, Bocconi University Press & Springer Verlag, 2011.

[132] G. Peccati, C.A. Tudor: *Gaussian limits for vector-valued multiple stochastic integrals*, *Séminaire de Probabilités XXXVIII*, p. 247-262, Lect. Notes in Math. vol. 1857, Springer Verlag, 2005.

[133] M. Plancherel, W. Rotach: *Sur les valeurs asymptotiques des polynomes d'Hermite* $H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n}\left(e^{-\frac{x^2}{2}}\right)$, Comment. Math. Helv., **1**(1929), 227-254.

[134] M. Reed, B. Simon: *Methods of Modern Mathematical Physics I. Functional Analysis*, Academic Press, 1980.

[135] S.O. Rice: *Mathematical analysis of random noise*, Bell System Tech. J., **24**(1945), 46-156.

[136] L. C. G. Rogers, D. Williams: *Diffusions, Markov Processes and Martingales. Volume 1. Foundations*, Cambridge University Press, 2000.

[137] N. Ross: *Fundamentals of Stein's method*, arXiv: 1109.1880.

[138] W. Rudin: *Fourier Analysis on Groups*, John Wiley & Sons, 1962.

[139] D. L. Saunders: *The Geometry of Jet Bundles*, LMS Series, vol. 142, Cambridge University Press, 1989.

[140] H. H. Schaefer with M. P. Wolf: *Topological Vector Spaces*, 2nd Edition, Springer Verlag, 1999.

[141] R. L. Schilling: *Measures, Integrals and Martingales*, Cambridge Univ. Press, 2005.

[142] R. L. Schilling, L. Partzch: *Brownian Motion. An Introduction to Stochastic Processes*, DeGruyter, 2012.

[143] D. Serre: *Matrices. Theory and applications*, 2nd Edition, Springer Verlag, 2010.

[144] I. Shigekawa: *Derivatives of Wiener functionals and absolute continuity of induced measures*, J.Math. Kyoto Univ., **20**(1980), 263-289.

[145] M. A. Shubin: *Pseudodifferential Operators and Spectral Theory*, 2nd Edition, Springer Verlag, 2001.

[146] B. Simon: *Convexity. An Analytic Viewpoint*, Cambridge University Press, 2011.

[147] M. Stecconi: *Kac-Rice formula for transverse intersections*, arXiv.2103,10853, Analysis and Mathematical Physics, 12(2):44, 2022.

[148] D. W. Stroock: *Gaussian Measures in Finite and Infinite Dimensions*, Springer Verlag, 2023.

[149] G. Szegö: *Orthogonal Polynomials*, Colloquium Publ., vol 23, Amer. Math. Soc., 2003.

[150] M. Talagrand: *Upper and Lower Bounds for Stochastic Processes. Decomposition Theorems*, 2nd Edition, Springer Verlag, 2021.

[151] M. Talagrand: *What Is Quantum Field Theory? A First Introduction for Mathematicians*, Cambridge University Press, 2022.

[152] A. Templeman: *Ergodic Theorems for Group Action. Informational and Thermodynamical Aspects*, Springer, 1992.

[153] V. A. Toponogov: *Differential Geometry of Curves and Surfaces. A Concise Guide*, Birkhäuser, 2006.

[154] F. Tréves: *Topological Vector Spaces, Distribution and Kernels*, Dover, 2006. (Reprint of the 1967 edition).

[155] N. N. Vakhaniya, V. I. Tarieladze: *Covariance operators of probability measures on locally convex spaces*, Teor. Veroyatnost. i Primenen., **23**(1978), 3-23). (in Russian); Engl. Transl Theor. Probab. Appl. **23**(1978), 1-21.

[156] V. A. Volkonskiĭ: *An ergodic theorem for the distribution of sojourn times*, Theor. Vorojastnost. i Primenen., **5**(1960), 357-360.

[157] H. Wenland: *Scattered Data Approximation*, Cambridge University Press, 2005.

[158] H. Whitney: *A function not constant on a connected set of critical points*, Duke. Math J. **34**(1935), 514-517.

[159] E. T. Whittaker, G. N. Watson: *A Course in Modern Analysis*, 4th Edition, Cambridge University Press, 1950.

[160] N. Wiener: *The ergodic theorem*, Duke Math. J., **5**(1939),1-18.

[161] A. M. Yaglom: *Correlation Theory of Stationary and Related Random Functions I. Basic Result*, Springer Verlag, 1987.

Index